11-9-2018

Generation of Isogenic Human Pluripotent Stem Cell-Derived Neurons to Establish a Molecular Angelman Syndrome Phenotype and to Study the UBE3A Protein Isoforms

Carissa Sirois
University of Connecticut - Storrs, csirois@uchc.edu

Generation of Isogenic Human Pluripotent Stem Cell-Derived Neurons to Establish a Molecular Angelman Syndrome Phenotype and to Study the *UBE3A* Protein Isoforms

Carissa L. Sirois, Ph.D.

University of Connecticut, 2018

Abstract

Angelman Syndrome (AS) is a neurodevelopment disorder for which there is currently no cure that is characterized by severe seizures, intellectual disability, absent speech, ataxia, and happy affect. Loss of expression from the maternally inherited copy of *UBE3A*, a gene regulated by genomic imprinting, causes AS. Currently there are multiple promising therapeutic approaches being explored and developed for AS, some of which involve targeting or expression of the human genetic sequence. Subsequently, it is necessary to establish robust cellular models for AS that can be used to test these, as well as future, potential AS therapies. Toward this aim, here we have used the CRISPR/Cas9 genome editing system to generate several isogenic human pluripotent stem cell lines two achieve two primary goals. First, we aimed to establish a robust quantitative molecular phenotype for cultured human AS neurons using the transcriptome. We identified and validated a list of genes that are consistent differentially expressed in AS neurons when compared to isogenic controls that can be assayed following drug treatments. Second, we aimed to study the abundance and localization of the three human UBE3A protein isoforms. We found that isoform 1 is the predominant protein isoform, and that UBE3A, regardless of isoform, appears to localize mostly to the cytoplasm, with low levels of expression in the nucleus and other organelles. The work in this thesis demonstrates that differentially expressed genes can be used as a phenotype for AS neurons to measure the effects of potential therapies, and provides important and previously unknown information as to the abundance and localization of the human UBE3A protein isoforms in human neurons.

Generation of Isogenic Human Pluripotent Stem Cell-Derived Neurons to Establish a Molecular Angelman Syndrome Phenotype and to Study the *UBE3A* Protein Isoforms

Carissa L. Sirois

B.S., Eastern Connecticut State University, 2010

M.S., University of Hartford, 2013

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2018

APPROVAL PAGE

Doctor of Philosophy Dissertation

Generation of Isogenic Human Pluripotent Stem Cell-Derived Neurons to Establish a Molecular Angelman Syndrome Phenotype and to Study the *UBE3A* Protein Isoforms

Presented by

Carissa L. Sirois, M.S.

Major Advisor _____
Stormy J. Chamberlain

Associate Advisor _____
Eric. S. Levine

Associate Advisor _____
Stephen J. Crocker

Associate Advisor _____
James Li

University of Connecticut
2018

Acknowledgements

I would first and foremost like to thank my advisor and mentor, Dr. Stormy Chamberlain. I am truly lucky to have had the opportunity to have Stormy has my mentor throughout my PhD because she was exactly the type of mentor that I needed. Stormy was always available to talk to about anything, whether it was my project, answering questions about various aspects of the PhD, or offering advice for after graduation. She always encouraged not only my work in the lab, but pursuing other endeavors such as applying for fellowships, giving talks, or attending conferences. She has greatly shaped the type of scientist that I am and that I strive to be.

Secondly, I would like to thank my thesis committee members: Dr. Eric Levine, Dr. Stephen Crocker, and Dr. James Li. My committee always supported the work that I did and provided great advice about my project throughout my PhD. Dr. Levine was very willing to collaborate with our lab on various aspects of my project, which was greatly appreciated. Dr. Crocker always had suggestions for follow up experiments at each of my committee meetings, as well as other helpful advice. Dr. Li rounded out my committee by providing great feedback on the genetic and developmental aspects of my project. I would also like to acknowledge Dr. Xue-jun Li, who served on my committee for part of my PhD before she left the university, and Dr. Marc Lalande, who I always considered to be my unofficial co-PI before he also left the university.

I have been incredibly lucky to work with a great group of people in the Chamberlain lab and, by close association, the Lalande lab and UConn Stem Cell Core. There are not enough pages in this thesis to fully describe here how much help and support I have received from Drs. Noelle Germain, Ivy Chen, and Jack Hsiao. These three individuals were the original members of the Chamberlain lab when I joined and they all taught me so much. I am grateful for all of their help and their continued friendship throughout my time here (and beyond!). I would also like to thank Chris Stoddard for always being willing to teach myself and other members of the lab, not only about genome editing, but also other various aspects of molecular biology. Thank

you also to the previous members of the Lalande lab, who were always such a joy to work with and who also taught me so much: Dr. Maeva Langouet, Dr. Souour Mansour, Dr. Elodie Mathieux, Erin Banda, Heather Glatt-Deely, and Michael Chung, who is now a member of the Chamberlain lab. I would also like to thank Dea Gorka, who worked with me as an undergraduate for two summers before joining our lab as a PhD student. Dea is a delight to work with and I hope that her and Michael continue the tradition of amazing students produced by our lab.

I would like to thank the faculty and students of both the Department of Neuroscience and Department of Genetics and Genome Sciences for any assistance and feedback that they provided during my time at UConn. A huge thank you to Dr. Justin Cotney, the members of the Cotney lab (Andrea Wilderman, Tara Yankee, and Jen Venoudenhove), and to Scott Adamson for all of their help with bioinformatics. Everyone was extremely patient with the many questions I had and was always willing to provide advice.

Thank you to all of the friends that I made during my time in the program: Judy Bloom, John Wizeman, Cory Brennick, Cory Willis, Alex Nicaise, Sarah Benjamin, Mitali Adlahka. This entire process was made so much better by sharing it with such an amazing group of people. Thank you also to all of my friends and family who have always been so supportive of my decision to be a perpetual student. A special thanks to my parents and sister, who have always encouraged me along the way and instilled in me the values of hard work, motivation, and dedication. Finally, I would like to thank my husband Sean, who has patiently waited ten years for me to no longer be a student. He has been the biggest source of support and I am so lucky to have him in my life.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 1

Angelman Syndrome, a genomic imprinting disorder of chromosome 15q11-q13.1

General introduction

The human chromosomal locus 15q11-q13.1 is subject to genomic imprinting[1] – the epigenetic phenomenon that produces a differential expression of a gene or genes in a parent of origin-specific manner. Unlike most genes which are expressed from both copies of a given chromosome, imprinted genes are expressed from only one copy of a chromosome – either the one inherited from the mother (the maternal allele) or the one inherited from the father (the paternal allele). In addition to imprinting, chromosome 15q11-q13.1 is also subject to chromosomal rearrangements (deletions or duplications) during meiosis due to the presence of regions of repetitive sequence (low-copy repeats) that are referred to as breakpoints.[2] During meiosis, these breakpoints can misalign and result in non-allelic homologous recombination, producing the subsequent deletions or duplications. Due to this combination of imprinted genes and susceptibility to deletions/duplications, three distinct disorders can arise from mutations in this region, depending on the nature of the mutation and which copy of the chromosome is affected.[3] Deletions of the paternal chromosome produce Prader-Willi Syndrome (PWS), deletions of the maternal chromosome produce Angelman Syndrome (AS), and duplications of the maternal chromosome cause 15q11-q13 Duplication Syndrome (Dup15q). Each of these disorders occurs at a frequency of about 1 in 15,000 live births.[4]

Genomic imprinting at chromosome 15q11-q13.1

Genomic Imprinting

While most genes in the mammalian genome are expressed biallelically, a small proportion of genes are subject to genomic imprinting. These genes carry epigenetic marks that are indicative of the chromosome's parent of origin (maternal or paternal), which results in the silencing of one copy of the gene. Most imprinted genes are found in clusters, and the imprinting

of each cluster is typically controlled by an imprinting control region (ICR). Imprinting occurs due to differential methylation at these ICRs, with the silenced allele being methylated and the active allele being unmethylated.[5,6] Methylation of the ICR occurs in the germline: during embryonic development, primordial germ cells migrate to the genital ridge, where they undergo global demethylation, causing the previous pattern of imprints from both parents to be erased. In oocytes, methylation of ICRs occurs after birth, while in male germ cells, methylation is mostly complete by birth. The majority of methylated ICRs are on the maternal allele.[5]

In addition to methylation at ICRs, imprinting of genes also can involve non-coding RNAs (such as *UBE3A-ATS, KCNQ1OT1*),[7] histone modifications (such as H3K4 methylation in male germ cells), and transcription.[6] While the exact evolutionary purpose of genomic imprinting is not known, it is known that imprinting in mammals is associated with evolution of the placenta. One theory is that imprinting could have evolved as a way to regulate the dosage of certain genes. Imprinted genes seem to be involved in important processes such as prenatal growth, placental growth, lineage development, normal brain function, and postnatal energy homeostasis. Mouse embryos in which both sets of chromosomes are inherited from one parent display gross developmental abnormalities.[5,8]

15q11-q13.1

The 15q11-q13.1 region is approximately 6 Mb in length, consisting of both imprinted and biallelically-expressed genes (**Figure 1**). The region contains 3 areas of low-copy repeats, known as breakpoints, that result in an increased likelihood for deletions and duplications to occur curing meiosis (**Figure 1, jagged lines**). The genes located between breakpoints 1 and 2 (BP1 and BP2) are biallelically expressed (*TUBGCP5, CYFIP1, NIPA2, NIPA1)*, while the region between BP2 and BP3 contains both imprinted (see below) and biallelically expressed genes (*GABRB3, GABRA5, GABRG3, OCA2, HERC2, GOLGA8*).

Imprinted Genes of 15q11-q13.1

The region of 15q11-q13 that is subject to genomic imprinting is approximately 2 Mb in length and contains both paternally-expressed and maternally-expressed genes.[3] Approximately 15 genes or transcripts are expressed exclusively from the paternal allele. *MKRN3, MAGEL2, NDN,* and *C15ORF2* are protein-coding genes. *SNURF-SNRPN* is a bicistronic transcript that encodes two proteins (*SNURF* and *SNRPN*) but also serves as host transcript for *SNORD107, SNORD64, SNORD108, SNORD109A, SNORD116, SNORD115, SNORD109B,* and the *UBE3A-ATS* transcripts.[9] This transcript spans approximately 600 kb of genomic DNA and undergoes alternative splicing. Additionally, the noncoding transcript *IPW* and two long noncoding RNAs (lncRNAs), *PWRN1* and *PWNR2,* are expressed from the paternal allele.[10] The SNORDs belong to a family of noncoding RNAs known as small nucleolar RNAs (snoRNAs, **Figure 1, vertical black lines**). In total, there are 72 snoRNAs produced from 15q11-q13, all of which are processed from introns of the *SNURF-SNRPN* transcript.[10] *IPW, SNORD115, SNORD 109B,* and *UBE3A-ATS* are expressed/produced specifically in neurons for reasons that are not fully understood. Imprinting of the paternally inherited genes is due to CpG methylation at *SNRPN, NDN,* and *MKRN3* on the maternal allele (**Figure 1, black circles**). While the silencing of genes on the maternal allele occurs in all cell types, there is neuron-specific imprinting of the *UBE3A* gene. In non-neuronal cells, *UBE3A* is expressed biallelically. Due to the neuron-specific production of the *UBE3A-ATS* transcript from the paternal allele (**Figure 1, dotted line**), however, the paternal copy of *UBE3A* is silenced in neurons.[11] The exact mechanism by which the *UBE3A-ATS* transcript silences the paternal copy of *UBE3A* is not known.

Bipartite imprinting center (AS-IC, PWS-IC)

Imprinting of the 15q11-q13 region is controlled by a bipartite imprinting center (IC), consisting of the Angelman Syndrome IC (AS-IC; **Figure 1, red shaded area**) and the Prader-

Willi Syndrome IC (PWS-IC; **Figure 1, blue shaded area**). These regions were initially identified as the shortest regions of overlap in patients with AS and PWS, respectively, with imprinting defect mutations.[12] The PWS-IC is approximately 4.3 kb in size and is comprised of the major promoter and exon 1 of the *SNURF-SNRPN* gene. Imprinting (silencing) of genes on the maternal allele requires the PWS-IC, as does maintenance of the imprint post-zygotically.[13,14] Additionally, the PWS-IC acts as the canonical promoter for *SNRPN* and the imprinted genes distal to it. The AS-IC is 35 kb upstream of the PWS-IC and is only 880 bp in length. While the exact mechanism is not clear, it is known that the AS-IC acts as a repressor of the PWS-IC so that the PWS-IC can be methylated on the maternal allele. Interestingly, while the AS-IC is required in gametes for the proper establishment of the imprints, it is not required later on in development.[15] Studies in mouse have confirmed that both the AS-IC and PWS-IC together are necessary for imprinting to occur,[14,15] and aberrant imprinting at these imprinting centers on either the maternal or paternal allele are sufficient to cause Angelman Syndrome or Prader-Willi Syndrome respectively.[12,16]

Disorders Arising from Mutations in 15q11-q13.1

The presence of low-copy repeats at certain parts of the chromosome make this region susceptible to deletions or duplications due to non-allelic homologous recombination during meiosis. Due to the imprinted expression of genes in this region, the phenotypes produced by the mutations are dependent upon which copy of the chromosome is affected.

Prader-Willi Syndrome

Loss of expression of the genes from the paternal copy of 15q11-q13 result in Prader-Willi Syndrome (PWS). Most (~70%) of PWS cases are caused by large deletions of the paternal 15q11-q13 region, with another 25% and 5% of cases being caused by maternal

uniparental disomy and imprinting defects of the paternal allele, respectively. PWS patients first present with hypotonia and failure to thrive in infancy, which manifests as decreased movement, lethargy, poor suck, and poor appetite. During childhood, patients develop increased appetite which progresses to hyperphagia, which in turn leads to morbid obesity. Additionally, patients present with short stature, small hands and feet, hypogonadism, motor and language delay, mild to moderate cognitive impairment, characteristic facial features, and behavioral phenotypes (temper tantrums, stubbornness, manipulative behavior, and obsessive-compulsive characteristics).[17] No single gene has been identified whose loss can account for the phenotype seen in PWS patients. By studying deletions of various sizes found in both typical- and atypical-deletion PWS patients, the smallest region identified that, when lost on the paternal allele, causes PWS has been narrowed down to approximately 91 kb of DNA that encompasses the *SNORD116* cluster and the *IPW* non-coding RNA.[18,19] How loss of these RNAs results in the PWS phenotype is not understood.

Angelman Syndrome

Loss of the maternal copy of 15q11-q13 results in a clinically distinct syndrome known as Angelman Syndrome (AS). AS is characterized by severe seizures, ataxia, motor delay, severe intellectual disability, absent speech, and happy demeanor. Patients also commonly present with microcephaly, sleep disorder, and hyperactivity. Patients typically begin presenting with some AS symptoms as early as 6 to 12 months of age, however the final diagnosis of AS typically doesn't occur until several years after the initial onset of symptoms. Infants with AS may have difficulty feeding, gastroesophageal reflux, and hypotonia. Other early (before 1 year of age) signs can include happy affect, microcephaly, strabismus, and tremulous movements. Seizures and specific changes in EEG tend to present by three years of age. Walking in AS patients typically does not occur until between 2 and 6 years of age, although 10% of AS patients are unable to walk.[20] AS is caused by loss of *UBE3A* from the maternal copy of the

chromosome.[21] Most AS cases (~75%) are caused by deletions encompassing 15q11-q13 on the maternal allele (**Figure 2A**). Mutations in the maternal copy of *UBE3A* account for another 10-15% of cases (**Figure 2B**), with paternal uniparental disomy (**Figure 2C**) or imprinting defects (**Figure 2D**) of the maternal allele accounting for the remaining cases.[22]

15q Duplication Syndrome

Duplications of the maternal 15q11-q13 region result in 15q Duplication Syndrome (Dup15q). Most (80%) Dup15q patients have a supernumerary isodicentric copy of chromosome 15, which typically contains two copies of the maternal 15q11-q13 region and two centromeres. As of result, these individuals have a tetrasomy of the 15q11-q13 region. The remaining Dup15q patients (20%) have an interstitial 15q11-q13 duplication, resulting in two copies of the region in tandem on their maternal copy of chromosome 15. Dup15q is characterized by hypotonia, motor delay, intellectual disability, epilepsy, and autism spectrum disorder. Indeed, maternal 15q11-13 duplication is the most common copy number variant associated with autism, accounting for 1-3% of all cases.[23] Like PWS, there is no one gene that accounts for the Dup15q phenotype. Because duplications of the paternal allele less severe phenotype that is not fully penetrant, it is assumed that *UBE3A* plays a major role Dup15q. However, duplications of maternal *UBE3A* alone are not sufficient to recapitulate Dup15q,[24] so it is likely that other genes within the duplication also contribute to the phenotype of these patients.

UBE3A

UBE3A protein function

AS is caused by loss of the maternal copy of 15q11-q13, which results in a loss of expression or function of maternal UBE3A. This is due to the fact that the paternal copy of *UBE3A* is imprinted in neurons, therefore AS-causing mutations affect the only functional copy

of *UBE3A* in these cells. The protein encoded by the *UBE3A* gene, also known as E6-associated protein (E6-AP), is an E3 ubiquitin ligase.[25] Its function is to assemble polyubiquitin chains on target proteins so that they can be degraded by the 26S proteasome. UBE3A is the founding member of the HECT (Homologous to E6-AP C Terminus) family of E3 ligases, whose ubiquitin ligase activity is conferred by a HECT domain.[26] Since there are AS patients with mutations in the region of *UBE3A* that encodes the HECT domain, resulting in truncated or ligase dead versions of the protein,[27,28] it is hypothesized that loss of UBE3A's E3 ligase activity causes AS.

Because of its role in AS and in certain cancers, there has been an enormous effort to identify robust UBE3A target proteins. Currently, only a handful of these targets are considered to be true and reliable substrates.[29–33] Unfortunately, the majority of these substrates are not necessarily relevant to AS or have not been demonstrated to be substrates in neurons. Many other proteins have also been identified as putative UBE3A substrates, however many putative substrates have not been validated as actual ubiquitin ligase targets.[34–36] In addition, a substantial portion of putative UBE3A substrates were identified in the context of various cancers.[37–39] For these reasons, it is currently unknown how loss of UBE3A in neurons specifically causes AS. In addition to its role as an E3 ubiquitin ligase, UBE3A has also been identified as transcriptional coactivator.[40] However, like the studies identifying its putative substrates, this role for UBE3A has yet to be demonstrated in neurons.

Furthermore, it has been demonstrated that UBE3A can be modulated by other proteins in certain contexts. For example, during human papilloma virus (HPV) infection, the HPV oncoprotein E6 acts as an allosteric activator of UBE3A, allowing it to target p53 for proteasomal degradation, which can lead to cervical cancer.[25,41] In the absence of E6, UBE3A normally does not ubiquitinate p53. Another example of a UBE3A modulator is another E3 ligase, HERC2, the gene for which is also located on chromosome 15q11-q13. HERC2 was shown to activate the E3 ligase activity of UBE3A *in vitro* and the two proteins were shown to

form a high molecular weight complex independent of the 26S proteasome.[35,42] Finally, in the process of trying to identify putative substrates, it has also been demonstrated that UBE3A is capable of interacting with other proteins without modulating their protein levels.[43–45]

Three protein isoforms of UBE3A

The human *UBE3A* gene produces multiple mRNAs that encode for three distinct protein isoforms.[46,47] All three human isoforms are full length versions of the UBE3A protein and presumably have E3 ligase function. Each protein isoform is translated from its own unique translational start site, however RNAs for isoforms 2 and 3 contain the translational start site for isoform 1.[48] Indeed, isoform 1 can be considered the "default" version of the UBE3A protein, with isoforms 2 and 3 having an additional 20 and 23 amino acids at their N-termini, respectively. The function of these additional amino acids is currently unknown. Additionally, it is not known whether the three isoforms share common substrates, localization, or if their abundance differs. While the three human isoforms were initially discovered more than 20 years ago, few studies have been conducted that explores their roles, abundance, or localization. One study of the three mouse protein isoforms demonstrated that dendritic outgrowth phenotypes in UBE3A-null cultured neurons could be rescued upon overexpression of mouse isoform 2, but not mouse isoforms 1 or 3.[49] This study also showed that the mouse isoforms tended to localize differently within the cell, with mouse isoform 3 localizing predominantly to the nucleus and isoforms 1 and 2 localizing to the cytoplasm in the soma and neuronal processes. The human and mouse protein isoforms, however, are not perfectly conserved. Mouse isoform 1 is a truncated version of the protein, lacking a HECT domain and therefore lacking E3 ligase activity. This isoform was shown to be related to miR-134 in mouse hippocampal neurons.[50] Additionally, human isoform 2 does not have a mouse equivalent.[48] Because of this, it is necessary to study the human isoforms of UBE3A to determine whether all three isoforms have a role not only in

normal biological function, but also in the various conditions where UBE3A is known to play a role (AS, Dup15q, HPV infection).

The prospect of molecular therapies for Angelman Syndrome

Symptom-based approaches to AS therapy

Currently AS is treated using a symptom management-based approach. The success of managing AS symptoms varies from patient to patient, and some symptoms (intellectual disability, absent speech) are currently not improved by available treatments.[20,51,52] Currently, anti-epileptic drugs (AEDs) are prescribed to control seizures, however no one drug is the gold standard for seizure treatment in AS. Additionally, it is common for seizure breakthrough to occur when patients are treated with a single medication, and some patients' seizures cannot be controlled. Behavior modification therapy can be useful for certain behaviors in AS that are considered to be socially disruptive or injurious, while other behaviors, such as hypermotoric behavior, tend not to improve with behavioral therapy. Other interventions shown to be beneficial in some AS patients include physical therapy, occupational therapy, speech therapy and orthopedic braces or surgery. Melatonin has been shown to improve sleep disturbances in some AS patients.[20]

At present, there have been only a few studies to examine whether specific drugs could improve any AS symptoms. A small study showed reduced seizures in four AS patients who were administered the corticosteroid prednisolone.[53] Another study of 23 AS patients showed improvements in seizures when patients were put on a low glycemic index diet, either alone or in conjunction with an AED.[54] An open label pilot clinical trial examining the effects of minocycline, a tetracycine antibiotic, on AS patients showed improvements on some subdomains of the outcome measures, such as communication (Bayley Scales of Infant and Toddler Development 3rd Edition), fine motor ability (BSID-III), and auditory comprehension (Preschool Language Scale 5th Edition).[55] A recent double-blind randomized placebo-controlled

study examined whether levodopa (L-dopa) administration lead to improvements patients, based on the findings that L-dopa reduced the excessive phosphorylation of CAMKII (Thr286 and Thr305/306) seen in AS mice. While well tolerated, L-dopa did not produce significant improvements in any of the measured outcomes.[56] Even more recently, it was announced by Ovid Therapeutics that gaboxadol, a delta-selective $GABA_A$ receptor agonist, produced significant improvements in AS patients in the clinical global impressions of improvement (CGI-I) scale when compared to placebo.[57]

Timing of UBE3A reinstatement in AS mice

A major consideration in the development of treatments for neurodevelopmental disorders is the earliest age at which patients can be treated, and whether intervention at that age, or later ages, will improve the quality of life for these patients. Angelman Syndrome is typically diagnosed during early childhood unless prior concerns resulted in prenatal or newborn genetic testing. There are two promising gene therapy approaches currently being explored for AS: unsilencing of the paternal copy of *UBE3A* or replacing UBE3A protein by delivery via adeno-associated virus (AAV)-based vectors. Both of these approaches rely on the assumption that restoring normal UBE3A levels in the brains of AS patients at a later developmental stage will be sufficient to overcome months or years of brain development in the absence of UBE3A. Recent studies in AS mice have indicated that there is indeed a critical period for UBE3A reinstatement to be effective at restoring disease-relevant phenotypes. Restoring UBE3A levels in AS mice embryonically rescued all phenotypes assayed (rotarod, marble burying, open field, nest building, forced swim, and seizure susceptibility), while partial restoration (30-50%) at P1-P5 rescued only rotarod and open field phenotypes. Interestingly, only rotarod phenotypes could be rescued by restoring UBE3A at juvenile (3 weeks) or adolescent (6 weeks) ages, while adult reinstatement (14 weeks) of UEB3A did not rescue any behavioral phenotypes.[58] Furthermore,

UBE3A reinstatement is capable of rescuing cellular phenotypes even at ages when there is no behavioral rescue: adult reinstatement of UBE3A was able to rescue deficits in long-term potentiation in the hippocampus[58] and synaptic transmission deficits in the medial prefrontal cortex.[59]

Treating AS by reactivation of the silent paternal UBE3A allele

The existence of a silenced but otherwise normal copy of the UBE3A gene on the paternal copy of chromosome 15 is one of the most promising therapeutic targets for AS. As stated above, the paternal copy of UBE3A is silenced in neurons due to the production of the *UBE3A-ATS* transcript. Gene therapies targeting this antisense transcript could therefore allow the paternal copy of UBE3A to be reactivated, and, if done early enough, could produce phenotypic improvement in AS patients. Evidence supporting this idea came from an AS mouse with a SV40 polyA and neomycin cassette inserted between *SNORD115* and *UBE3A*, which was designed to prevent the *UBE3A-ATS* from overlapping the *UBE3A* gene. This approach, when used in both wildtype and AS mice, was able to upregulate *UBE3A* expression from the paternal allele. It also improved marble burying phenotypes, wire hanging test performance, dowel test performance, contextual fear conditioning, and long-term potentiation.[60]

One early attempt to achieve paternal *UBE3A* unsilencing in AS patients was by administering “promethylation supplements”, which would theoretically increase global DNA methylation.[61,62] It was hoped that this global methylation would include methylation of the PWS-IC on the paternal allele, which would silence SNURF/SNRPN and therefore prevent production of the *UBE3A-ATS*. Two studies examined the effects of these supplements on AS patients. The first was a double-blind placebo-controlled study of the effects of betaine and folic acid.[63] This study did not demonstrate any effects of the treatment on levels of global methylation (as measured by levels of methionine and homocysteine) or on developmental outcome measures. A second study, a one-year open label study, examined the combination of L-5-

methyltetrahydrofolate, vitamin B12, betaine, and creatine. As with the first study, there were no changes in global DNA methylation or on behavioral outcome measures.[64]

More recent and promising approaches have focused on targeting the already existing lncRNA instead of preventing its production. A high-content small molecule screen found that topoisomerase type I (TOP1) and type II (TOP2) inhibitors could unsilence the paternal copy of UBE3A in cultured mouse neurons. One such TOP1 inhibitor, topotecan, was the most potent compound and already an FDA-approved drug for cancer. It was demonstrated that topotecan unsilenced paternal UBE3A by causing a downregulation of the *UBE3A-ATS,* both *in vitro* and *in vivo*.[65] It was also shown that the formation of TOP1 cleavage complexes upon topotecan treatment are required to unsilenced paternal *UBE3A.*[66] Further studies determined that topotecan decreases the expression of long genes in both mouse and human neurons, including almost all genes greater than 200 kb in length, and that many of these long genes are autism candidate genes and/or genes that encode synaptic proteins, including *NRXN1, NLGN1, CNTNAP2,* and *GABRB3*.[67] Subsequently, it was shown that topotecan treatment suppresses spontaneous network activity and inhibitory neurotransmission, while also reducing excitatory neurotransmission.[68] Due to a limited bioavailability in the CNS and reports of toxicity, alternative TOP1 inhibitors, specifically those derived from indenoisoquinoline, have since been investigated. All thirteen of the indenoisoquinoline-derived TOP1 inhibitors tested were able to unsilenced the paternal copy of UBE3A in cultured mouse neurons.[69] Two of these compounds, indotecan and indimitecan are currently in clinical trials for individuals with solid tumors and lymphoma.

A second approach used to target the *UBE3A-ATS* transcript was to design antisense oligonucleotides (ASOs) specific to the *UBE3A-ATS*. ASOs are short segments of chemically-modified oligonucleotides that are designed to be complementary to a transcript of interest. The effect of the ASO on the target transcript varies depending on the design of the oligo. ASOs with a “gapmer” design (a short segment of unmodified nucleotides flanked by nucleotides with a

modified sugar ring) work by inducing RNase H-mediated degradation of their target upon binding. This approach was taken to target the *UBE3A-ATS in vitro* and *in vivo*.[70] *In vivo* administration of the ASOs via intracerebroventricular (i.c.v.) injection produced long-term unsilencing (16 weeks) of the paternal allele following a single injection. Additionally, adult mice treated with ASO showed improvements in contextual freezing and obesity phenotypes seen in AS mice.[70]

Expression of exogenous UBE3A using adeno-associated virus-based approaches

Another therapeutic approach being pursued is to replace UBE3A in AS patients by introducing a normal copy UBE3A via an adeno-associated virus (AAV)-based vector. This approach was demonstrated to be beneficial in a mouse model of AS using a recombinant AAV.[71] AS mice receiving the UBE3A-AAV expressed normal levels of UBE3A protein, showed a recovery in early phase LTP and post-tetanic potentiation, and showed improvements in both contextual fear conditioning and the Morris water maze. Currently, an AAV-based approach for UBE3A replacement is in the preclinical stages of development.[72]

The work outlined in this thesis details the generation and use of isogenic human pluripotent stem cell models to study Angelman Syndrome and UBE3A. Generation of these cell lines is outlined in **Chapter 2**. These stem cell lines were then used to establish a robust quantitative molecular phenotype for AS (**Chapter 3**) or to study the three human UBE3A protein isoforms (**Chapter 4**). Knowledge gained from the use of these stem cell models will contribute information that is necessary for the development and testing of safe and effective AS therapies in cultured human neurons.

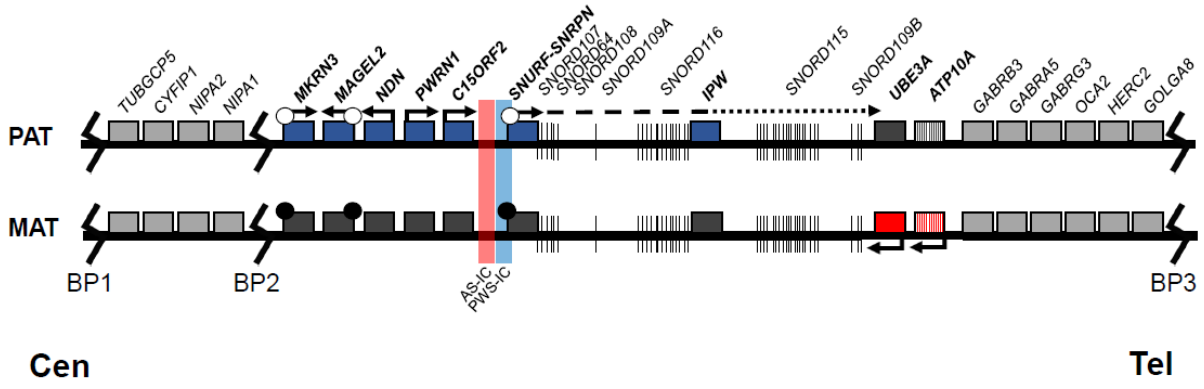


Figure 1. Map of chromosome 15q11-q13.1 between BP1 and BP3

Paternally-expressed genes are shown in blue. Maternally expressed genes are shown in red. Biallelically expressed genes are shown in light grey. Silenced genes are shown in dark gray. Circles indicate differentially methylated regions. Open circles are unmethylated, closed circles are methylated. ATP10 is shown as striped because its imprinting status varies between individuals. Arrows indicate the direction of transcription. Jagged lines indicate the breakpoints caused by low-copy repeats. The dashed/dotted line indicates the long non-coding transcript produced by SNURF/SNRPN that serves as the host transcript for the *UBE3A-ATS* as well as the snoRNAs. (Figure based on Chamberlain & Lalande, 2010)[3]

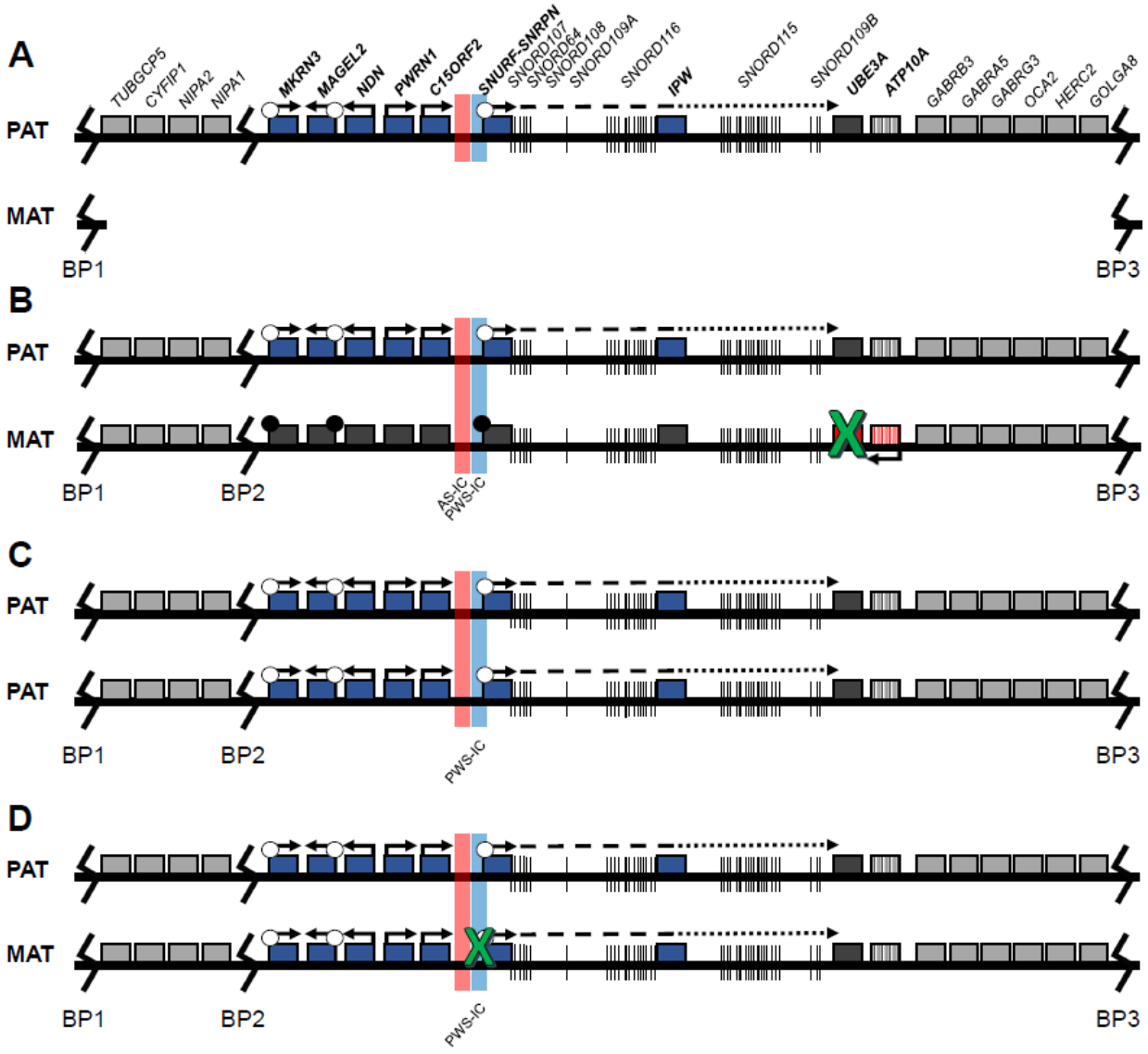


Figure 2. Genetic causes of Angelman Syndrome

A Deletion of maternal 15q11-q13.1 between BP 1 and BP 3
B Mutation in maternal *UBE3A*
C Paternal uniparental disomy (UPD)
D Imprinting defect of maternal allele

Chapter 2

Generation of isogenic human pluripotent stem cell lines using CRISPR/Cas9

Some data presented in this chapter was published in the following paper:

Title: Disrupted neuronal maturation in Angelman syndrome-derived induced pluripotent stem cells

Authors: James J. Fink[1], Tiwanna M. Robinson[1], Noelle D. Germain[2], Carissa L. Sirois[2], Kaitlyn A. Bolduc[2], Amanda J. Ward[3], Frank Rigo[3], Stormy J. Chamberlain[3], & Eric S. Levine[1]

[1] Department of Neuroscience, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, Connecticut 06030, USA.

[2] Department of Genetics and Genome Sciences, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, Connecticut 06030, USA.

[3] Ionis Pharmaceuticals, Carlsbad, California 92010, USA.

My contribution: generation & characterization of UBE3A KO line

Background & Rationale

Each person's genome contains its own unique combination of copy number variants (CNVs) and single nucleotide variants (SNVs), which can affect overall mRNA expression levels and mRNA splicing.[73,74] This, coupled with epigenetic factors that can also influence gene expression, results in large-scale differences in gene expression between unrelated individuals. These gene expression changes can introduce significant variation when trying to establish robust cellular models and phenotypes in an *in vitro* system. Isogenic cell lines remove this complication and can also eliminate the need to use a large number of samples, as is typically the case when comparing cells from unrelated individuals. The recent adaptation of the bacterial adaptive immune system CRISPR-Cas9 for use as an RNA-guided genome editing tool has greatly improved the ease with which we can edit the genome *in vitro*, therefore allowing us to more easily generate isogenic cell lines. Importantly, we can edit induced pluripotent or human embryonic stem cells using CRISPR technology, allowing us to make precise changes to the genome in human cells, which can then be differentiated into disease-relevant cell types.

Angelman Syndrome (AS) is caused by the loss of expression of the *UBE3A* gene on the maternal copy of chromosome 15q11-q13.[21] Currently there is no treatment or cure for AS, however multiple lines of evidence suggest that AS is an ideal candidate for gene therapy-based approaches. First, studies in mouse have indicated that restoration of UBE3A protein levels postnatally can have a phenotypic benefit, depending on the timing of the intervention.[58,70] AS mice in which Ube3a expression has been rescued at a juvenile age recover motor function deficits seen during the rotarod task, show some improvement in the forced swim and marble burying tasks, and have rescued long-term potentiation (LTP), while mice with adolescent or adult Ube3a restoration show less phenotypic improvement. This indicates that therapeutic intervention in humans, which in most cases wouldn't occur until early childhood, may still have some therapeutic benefit. Second, the existence of a perfectly good copy of the UBE3A gene on

the silenced paternal allele is a promising therapeutic avenue to explore. Indeed, it has been shown in mice and in human stem cell derived neurons that preventing the *UBE3A-ATS* transcript from silencing *UBE3A*, either by degrading it or by preventing its production, results in reactivation of the silent paternal copy of *UBE3A*.[60,65,67] Furthermore, it has been shown that antisense oligonucleotides specific to the *Ube3a-ats* in mouse can restore paternal *Ube3a in vivo* and restore some phenotypes in AS mice treated at 2 to 4 months of age.[70]

While the work done in mice has been promising, further advances in developing cures or treatments for AS will require the use of human cells or, at the very least, knowledge of the human genetic and protein sequences for *UBE3A* and *UBE3A-ATS* in order to develop therapies that are effective in human patients. Therapeutic approaches that involve targeting the *UBE3A-ATS* transcript need to be tested in human cells, or humanized rodent models, as the human transcript is not identical to the mouse transcript.[47,75] Vector-based gene therapy approaches[72] that involve delivery of a UBE3A transgene(s) will require use of the human cDNA sequence and also require knowledge of the human UBE3A protein isoforms to know which versions of the protein will need to be replaced in patients. Finally, it will also be important to test the toxicity of any and all potential AS therapies in human neurons.

This chapter outlines the use of the CRISPR/Cas9 genome editing system to edit the *UBE3A* gene in induced pluripotent stem cell (iPSC) and human embryonic stem cell (hESC) lines. Some of these cell lines have been used to establish (**Chapter 3**) or confirm[76] robust cellular phenotypes for AS stem cell-derived neurons, allowing us to examine these phenotypes in cells that are genetically identical aside from their respective *UBE3A* mutations and be confident that these phenotypes are not an artifact of person-to-person variability. Other lines generated here have been used to study the abundance and localization of the UBE3A protein isoforms (**Chapter 4**). For this project it was important to make isogenic stem cell lines for two reasons. First, we are only aware of AS patients with mutations affecting one of the three isoforms (isoform 1[77]), which prevents us from being able to generate iPSCs with mutations

affecting the other isoforms simply by reprogramming AS patient fibroblasts. Second, since the goal was to examine total UBE3A protein levels upon loss of individual isoforms, any subtle changes in protein abundance or localization could be lost by comparing non-isogenic cell lines. The knowledge gained by studying these isogenic stem cells and neurons will be useful for the development of potential AS therapies: the phenotypes that have been established in isogenic AS cells can be assayed following drug treatments, while knowledge of the abundance of each isoform can be used to inform UBE3A replacement therapies.

Materials and Methods

iPSC/hESC culture

Genome editing was performed in AS Pt Mut iPSCs, MCH2-10 iPSCs, and H9 (WA09) hESCs. AS Pt Mut iPSCs were derived by reprogramming fibroblasts from a male AS patient with a point mutation (T→C) in the *UBE3A* gene, which results in an amino acid change (F583S). This mutation was inherited by both the proband and his brother (also diagnosed with AS) from their mother, who in turn had inherited it from her father. MCH2-10 iPSCs were derived from a neurotypical female and have been previously described.[78,79] All stem cells were cultured and maintained as previously described.[78,79] Briefly, stem cells were grown on irradiated mouse embryonic fibroblasts (MEFs), fed daily with iPSC/hESC media, and manually passaged every six to seven days. iPSC/hESC media consisted of DMEM/F12 supplemented with knock-out serum replacement (KOSR), L-glutamine plus β-mercaptoethanol, non-essential amino acids (NEAA), and 2x basic FGF. All cells were grown in a humid incubator kept at 37°C and 5% CO_2.

Materials for CRISPR/Cas9-mediated genome editing

The sgRNA used to edit the Rx35i-7 iPSC line and the sgRNA located at the 3' end of *UBE3A* used to delete *UBE3A* in H9 hESCs were both designed by Chris Stoddard of the hESC/iPSC Targeting Core at the University of Connecticut Health Center. The sgRNA used in the Rx35i-7 iPSCs was designed to preferentially target the maternal *UBE3A* allele, which contained the AS-causing point mutation in these iPSCs. The core also designed and constructed a targeting vector used in the initial attempts to edit this cell line (**Figure 1A**). The 3' UBE3A gRNA targets the last exon of the gene, just upstream of the STOP codon (**Figure 5**). All other sgRNAs and all single-stranded oligonucleotides (ssODNs) were designed by the author. These sgRNAs were designed using MIT's CRISPR design website

(http://crispr.mit.edu). Three sgRNAs were designed to target the translational start sites of the UBE3A protein isoforms, with one sgRNA designed per isoform/start site (**Figure 4**).

All sgRNAs were cloned into the plentiCRISPRv2[80] (AS Pt Mut iPSCs), pX330[81] (AS Pt Mut iPSCs; UBE3A deletion 3' CRISPR), or pX459v2[82] (all other lines) vectors, which also contain human codon-optimized SpCas9 and puromycin resistance cassette. Design of ssODNs was based on work from Yang et al (2013).[83] Each ssODN contained the intended correction or mutation of the *UBE3A* gene, as well as a silent mutation to create a restriction digest site. This site was introduced to facilitate the screening of a large number of clones, as well as to disrupt the PAM sequence, which would prevent Cas9 from re-cutting an edited allele. For all silent mutations, the codons used were confirmed to be ones already used elsewhere in the gene. This ensured that the cell already had the appropriate transfer RNA (tRNA) for that codon, which could then be used during translation. The ssODNs used to generate the Isoform 2 and Isoform 3 KO lines were ordered as the reverse compliment to the strand containing the PAM sequence and the 5' ends phosphorylated. All sgRNAs and ssODNs are shown in **Tables 1 & 2**.

Electroporation or Nucleofection of iPSCs/hESCs

Stem cells were incubated in 10 μM ROCK inhibitor (ROCKi, Y-27632) for 6 to 24 hours prior to electroporation. Cells were singlized using Accutase, pelleted by centrifugation, then resuspended in cold PBS. Resuspended cells were loaded into a cuvette containing 10 μg CRISPR vector and, when appropriate, 2-6 μl of 100 μM ssODN, then electroporated using the Gene Pulser X Cell (BioRad). Electroporated cells were centrifuged to remove cell debris then plated onto a 10 cm DR4-MEF dish in media containing ROCKi and the small molecule L755507 (5 μM; Xcessbio), which was reported to enhance homology directed repair (HDR)[84]. For some experiments, when relevant, the small molecule SCR7 was also added to the media as it had been reported to block NHEJ.[85] Puromycin selection was started 24 hours after electroporation and maintained for 48 hours. L755507, SCR7 and ROCKi were removed from

the media 48 hours after electroporation. Cells were then maintained in regular iPSC/hESC media until large enough to screen, at which point they were put into ROCKi and penicillin/streptomycin overnight prior to screening. Electroporation was used to generate the following cell lines: correction of AS Pt Mut iPSCs (Corrected iPSC line), knockout of UBE3A in MCH2-10 iPSCs (MCH2-10ΔU line), and mutation of the isoform 1 translational start site in H9 hESCs (Iso1KO line). For correction of the AS Pt Mut line, early attempts to correct the UBE3A mutation were done by electroporating the CRISPR (10 µg) and a targeting vector (40 µg) containing a floxed neomycin resistance cassette. These cells were selected with G418 (50 µg/mL) starting 72 hours after electroporation and maintained in G418 for 5 days.

Cells edited by nucleofection were treated similarly to electroporated cells with the following changes: singlized stem cells were resuspended in a solution containing nucleofector solution, supplement solution, 2 µg of CRISPR vector and 3 µl of 100 µM ssODN (Iso2KO and Iso3 KO lines), or 2.5 µg of each CRISPR vector (ΔUBE3A$^{m-/p+}$ and UBE3A KO lines). Resuspended cells were loaded into a cuvette and nucleofected using the Amaxa 4D Nucleofector (Lonza). Nucleofection was used to generate the following cell lines: mutation of isoform 2 translational start site (Iso2KO line), mutation of isoform 3 translational start site (Iso3 KO line), and deletion of UBE3A in H9s (ΔUBE3A$^{m-/p+}$ and UBE3A KO lines). For each electroporation, 6-8 x 10^6 cells (3 wells of a 6 well plate) were used, while for each nucleofection 2-3x10^6 cells (1 well of a 6 well plate) were used.

Clone Screening

Individual colonies were cut into pieces using sterile needles, then half of the colony transferred into a well of a round-bottom 96 well plate containing stem cell media, ROCKi, and pen/strep. The other half of each colony was harvested for HotSHOT (hot sodium hydroxide and tris) genomic DNA (gDNA) preparation. Cells were pelleted by spinning in a tabletop centrifuge for 30 to 60 seconds. Cell culture medium was removed then the pellet incubated in 25 µl of

alkaline lysis reagent (25 mM NaOH, 0.2 mM EDTA) for 45 minutes at 95°C. Next 25 µl of neutralization reagent (40 mM Tris-HCl) was added, then the tubes were mixed by vortexing (2-3 brief pulses). Resulting HotSHOT gDNA was immediately used in conventional PCR reactions specific to each genome editing experiment (see Results section). At the end of the clone screening process, positive clones were transferred from round bottom wells into individual wells of a 12 or 24 well MEF plate and kept in ROCKi and pen/strep overnight.

Conventional PCR

Conventional PCR reactions were performed using either Herculase II Fusion DNA Polymerase (Agilent) or Platinum Taq DNA Polymerase High Fidelity (ThermoFisher). Primers used for each genome editing experiment are listed in **Table 4**.

Stand-specific RT-PCR

Strand-specific RT for the UBE3A-ATS transcript was performed on 1 µg of DNase I-treated RNA isolated from 10-week-old H9 hESC-derived neurons using Superscript III Reverse Transcriptase (ThermoFisher Scientific). RT was performed according to manufacturer's instructions with the following exception: the 60-minute incubation was done at 53°C instead of 55°C, as recommended for gene-specific primers. RT-PCR was performed on the stand-specific cDNA libraries using the Advantage 2 Polymerase PCR Kit (Takara Bio). PCR reactions were prepared using the 10X Advantage SA Buffer and run on a thermal cycler set at 94°C/3min; [94°C/15sec, 60°C/30sec, 68°C/30sec] x 32 cycles; 72°C/10min. Primers used for RT and RT-PCR are listed in **Table 4**.

Western Blot

Western blot to confirm knockdown of UBE3A in the MCH2-10ΔU line was performed as described in Fink et al (2017).[76] Briefly, cells were lysed using PathScan Lysis Buffer (Cell Signaling Technology) supplemented with 1 mM phenylmethane sulfonyl fluoride (PMSF; Gibco)

and Protease Inhibitor Cocktail III (used at 1:200; Calbiochem (EMD Millipore)). Fourteen micrograms of total cell lysate were separated by SDS-PAGE using 4-20% TGX Stain-Free mini gels (BioRad). Protein was transferred to PVDF membranes using the TransBlot Turbo system (BioRad). Membranes were blocked in 5% milk in TBS-T (Tris-buffered saline plus 0.1% Tween-20) for 1 hour at room temperature then incubated in blocking buffer containing primary antibodies overnight at 4°C (UBE3A) or at room temperature for an hour (GAPDH). Membranes were washed with TBS-T at room temperature, incubated in blocking buffer containing HRP-conjugated secondary antibodies for 1 hour at room temperature, then washed again in TBS-T. Membranes were imaged using the Clarity Western ECL substrate (BioRad) on the ChemiDoc Touch Imaging System (BioRad). The following primary antibodies were used: rabbit anti-UBE3A (Bethyl Laboratories Inc., 1:1000), mouse anti-GAPDH (EMD Millipore, 1:10,000). Secondary antibodies (Cell Signaling Technologies) were used at the following concentrations: anti-rabbit-HRP 1:3000, anti-mouse-HRP 1:10,000

Results

Correction of UBE3A point mutation in AS iPSCs

A. Correction of mutation using CRISPR + targeting vector

AS Pt Mut iPSCs were electroporated with pX330 vector containing a sgRNA specific to the mutant allele and targeting vector containing the wild type *UBE3A* sequence to use as a template for HDR. Four clones survived 5 days of selection with G418 and no additional colonies appeared after switching the plate to regular media for an additional 8 days. The four clones were expanded and genomic DNA isolated. Clones were screened by conventional PCR with primers within and outside of the targeting vector homology arms (**Figure 1B** and **Table 4**). One clone (#1) was positive and sent out for Sanger sequencing to determine whether the mutation was corrected. Sanger sequencing revealed correction of the mutation in this line and the presence of a silent mutation introduced to identify the edited allele in downstream applications (**Figure 1C**). Clone 1 was then nucleofected with CRE-IRES-Puro plasmid to remove the floxed Neo cassette that should have been inserted during the electroporation. All clones screened did not have any loxP sites present following nucleofection even though one loxP would be expected to remain following CRE-mediate recombination (**Figure 1D**). It is likely that the 5' arm of the targeting vector was not integrated during the genome editing process and only the 3' end (downstream of the Neo cassette) was used as a template for HDR. Ultimately, this cell line was not used for any further work as it was determined that this stem cell line was not the one derived from the AS patient but from the patient's mother.

B. Correction of mutation using CRISPR + ssODN

AS Pt Mut iPSCs were corrected to a wildtype genotype by electroporation of cells with plentiCRISPRv2 vector containing the same sgRNA as above and an ssODN. This ssODN contained both the wildtype sequence for the region of interest in the UBE3A gene (to

correct the AS mutation) and a silent mutation that created a FokI restriction site (**Figure 2A** and **Table 2**). Sixty-four clones were screened by conventional PCR of HotSHOT gDNA, followed by FokI digest of the PCR product. Clones with a positive digest were expanded then their genotype confirmed by Sanger sequencing of PCR products off of phenol chloroform-extracted gDNA to determine whether the AS mutation had been corrected. The first electroporation did not yield any positive clones out of the 40 screened, while the second electroporation yielded one clone (#22) out of 64 that had digested and that, upon sequencing, had both the silent mutation and correction of the point mutation (Corrected line; **Figure 2C**). Both the AS Pt Mut and Corrected iPSC lines were used for the work detailed in **Chapter 3** of this thesis.

Mutate translational start sites for UBE3A protein isoforms in H9 hESCs

Each UBE3A protein isoform is translated from its own unique translational start site, however the methionine used as the translational start site for Isoform 1 is present in the other two UBE3A isoforms (**Figure 3A**).[47,48] Using CRISPR to create an indel at this location would therefore disrupt the reading frame of all three protein isoforms. For this reason, we chose to change the translational start sites of each protein isoform from a methionine to a leucine, which would prevent the methionine from being used as a start site. Replacing the methionine that serves as the Isoform 1 translational start site with a leucine was predicted to not cause any deleterious changes to the structure or folding of the remaining two UBE3A protein isoforms when analyzed with multiple protein prediction softwares (**Figure 3B-C**).[86,87] The sgRNAs and ssODNs used to generate the three isoform knockout hESC lines are shown in **Tables 1 & 2**. The ssODNs contained sequences to change the appropriate methionine codon (ATG) into a leucine codon, as well as a silent mutation to create a restriction digest site and disrupt the PAM, which would prevent cutting of the edited allele(s). Clones were screened by conventional PCR of HotSHOT gDNA followed by restriction digest of the PCR product with TaqI (Isoform 1

KO), HindIII (Isoform 2 KO), or HPYCH4V (Isoform 3 KO). Digest positive clones were then sequenced by Sanger sequencing. Using this method, we were able to identify one homozygous clone for each of the three isoform start sites in which the methionine codon had been changed to a leucine codon, resulting in three edited hESC lines (Clone #15 (Iso1), #60 (Iso2), and #77 (Iso3); **Figure 4B-D**). To obtain these lines, two, four, and four rounds of electroporation or nucleofection were performed to mutate Isoform 1, 2, and 3, respectively. The Isoform 1 KO clone was originally a mixed clone with some cells possessing the wildtype sequence and silent mutation but the majority of the cells possessing only the start site mutation. Upon routine passaging to maintain the cell line, the small population of silent mutation + wildtype start site cells was eventually removed (**Figure 4B**). The isoform KO lines were used in work detailed in **Chapter 4** of this thesis.

Deletion of UBE3A in H9 hESCs

To generate an additional pair of isogenic AS and control stem cells, we deleted the maternal copy of *UBE3A* in H9 hESCs. To do this, we employed the strategy described by Kraft and colleagues, where a pair of CRISPRs flanking a region of interest was used to generate clones with deletions, duplications, and/or inversions of that region.[88] To delete *UBE3A*, we used the isoform 1 translational start site CRISPR described above (located at the 5' end of the gene) and a CRISPR located at the 3' end of the gene just upstream of the STOP codon, which had been previously used by our lab to insert a GFP into the UBE3A gene (**Figure 5A**). Clones were screened for the presence of deletions, duplications, or inversions using different combinations of primers (**Figure 5A-C** and **Table 3**). Clones identified by PCR as being heterozygous for a deletion of *UBE3A* were Sanger sequenced to determine whether were any indels on the intact allele (**Figure 5D**), as both CRISPRs are located in exons. Using this method, we were able to identify one clone with a maternal deletion of UBE3A (#28; $\Delta UBE3A^{m-/p+}$) and multiple clones with both alleles of UBE3A either deleted or mutated, resulting in

homozygous knockouts (UBE3A KO). While homozygous KO clones were identified during every round of genome editing, no clean heterozygous clones were identified until the 5th round of genome editing. The H9ΔU$^{m-/p+}$ line was used in work detailed in **Chapter 3** of this thesis, while one of the UBE3A KO clones was used as a control line for experiments detailed in **Chapter 4**.

To determine the identity of the deleted allele in the heterozygous deletion clone, we took advantage of the presence of SNPs near the 5' CRISPR cut site (**Figure 6A**). Sanger sequencing of H9 hESC gDNA confirmed the presence of 4 SNPs within 3.1 kb of this location (**Figure 6B**). Strand-specific RT-PCR (ssRT-PCR) was performed to determine the allelic identity of each SNP. To do this, we used RNA from H9 10-week neurons in order to sequence the *UBE3A-ATS*, as it is produced exclusively from the paternal allele in neurons and is complimentary to the gDNA of the paternal copy of the gene. In order to sequence these specific SNPs, primers complementary to the *UBE3A-ATS* at regions near the SNPs of interest were used in the ssRT reaction. ssRT-PCR products were then Sanger sequenced to identify the SNPs (**Figure 6C**). Because rs1041933 had the cleanest sequencing trace, we used that SNP to determine which allele was deleted in our heterozygous clone. Conventional PCR was first performed to amplify the intact allele, then a nested PCR using primers located near the SNP were used on the gel-purified product of the first reaction. Finally, the second product was sent for Sanger sequencing. The presence of a G at this SNP indicated that the intact allele was the paternal copy of the chromosome (**Figure 6D**) and therefore the maternal allele was deleted.

Knockout of UBE3A in normal iPSCs

The Isoform 1 start site CRISPR was also used to generate a UBE3A knockout line in MCH2-10 normal iPSCs. iPSCs were electroporated with CRISPR plasmid in hopes that cutting by Cas9 would result in non-homologous end joining (NHEJ) and produce an indel that would

shift the reading frame for the protein. Since the methionine used as the isoform 1 start site is present in all three protein isoforms, an indel at this location would likely disrupt all three protein isoforms. The goal was to mutate both copies of UBE3A so as to avoid having to screen for clones with only the maternal copy of UBE3A mutated. Clones were screened by Sanger sequencing of the CRISPR cut site to check for indels. One clone (#33) had a one base pair insertion (**Figure 7A**) but appeared to be a mixed clone as the sequencing trace had low levels of wildtype sequence. TIDE analysis (Tracking of Indels by DEcomposition[89]) confirmed that at least 12% of the sample was still wildtype. This clone was singlized and subclones were resequenced and reanalyzed by TIDE, confirming the absence of wildtype gDNA (**Figure 7B**). Knockdown of UBE3A protein in the iPSCs was confirmed by Western blot (**Figure 7C**). These stem cells were differentiated into neurons and used to determine whether the UBE3A KO neurons recapitulated electrophysiological phenotypes seen in AS iPSC-derived neurons when compared to their isogenic control.[76]

Discussion

This chapter demonstrates the successful generation of isogenic stem cell lines using CRISPR/Cas9-mediated genome editing. Generation of these stem cell lines was necessary for the experiments outlined in the next two chapters and have also been critical components of experiments done in the labs of collaborators to establish robust cellular phenotypes in AS neurons. As discussed in **Chapter** 3, the AS Pt Mut/Corrected and H9/ΔUBE3A$^{m-/p+}$ lines were used to establish a molecular transcriptome phenotype for AS stem cell derived neurons, while in **Chapter** 4, the Isoform KO lines were used to examine the abundance and localization of the UBE3A isoforms. The MCH2-10ΔU iPSCs was used to confirm that electrophysiological phenotypes seen in AS iPSC-derived neurons were a direct result of UBE3A deficiency.[76] These iPSCS, as well as the AS Pt Mut/Corrected iPSCs, and ΔUBE3A$^{m-/p+}$ hESC lines are currently being used to study the morphology of cultured AS stem cell-derived neurons. Furthermore, the AS Pt Mut/Corrected and MCH2-10/MCH2-10ΔU lines are available for distribution to other institutions, where they can be used to ask other questions relevant to AS research. In addition to the above applications, there are many other potential applications of these stem cell lines. As seen in **Chapter 4**, the UBE3AKO line makes an excellent negative control for antibody-based applications such as immunocytochemistry, Western blot, or ChIP-seq. Having isogenic AS and control pairs will be useful in establishing other phenotypes for stem cell-derived AS neurons, but also could be serve as a negative control in proteomics approaches to identify putative neuronal targets of UBE3A.

As demonstrated above, genome editing experiments that rely on HDR to make specific changes to the genome are relatively variable in their efficiency. As CRISPR is still a relatively new technology, the field is constantly changing and improving the methods used to conduct these experiments. Indeed, during the course of the work described above, multiple changes were made to the way we designed our oligos and performed genome editing experiments. For

example, we found that delivery of the CRISPR plasmid and ssODN via nucleofection appeared to improve the number of positive clones identified during clone screening. The use of ssODNs as a template for HDR was also a major change, since previously all work had been done using targeting vectors, as in **Figure 1**. We also began to phosphorylate the 5' ends of the ssODNs. Finally, multiple small molecules were published that purported to either enhance HDR (such as L755507[84]) or inhibit NHEJ (such as SCR7[85]), both of which were used when appropriate.

Even with these improvements, however, these experiments were still relatively inefficient, especially the mutations of the Isoform 2 and 3 start sites. Future attempts at HDR could make one or more of the following changes in an attempt to further improve efficiency. To generate cell lines with heterozygous corrections or mutations without indels on the other allele, one approach is to use two ssODNs containing a silent mutation – one with the WT sequence and one with the intended edit. In this case, Cas9 has generated a double-stranded break on both alleles and both ssODNs are used as templates for HDR.[90] During the process of screening clones, we encountered multiple instances of putatively positive clones that digested during screening that would also have indels, which were only revealed upon Sanger sequencing. Positive clones identified using the double ssODN method would be less likely to have indels on the wildtype allele. In the case of the isoform mutation cell lines above, if it had been necessary to have heterozygous start site mutations on the maternal allele, this method could have been employed to decrease the likelihood of finding clones with indels. While we were fortunate to be able to use ssODNs as templates for HDR to generate our final cell lines, there are certain situations that necessitate the use of a targeting vector as the template for HDR. For example, fluorescent proteins are too large to insert using an ssODN. Recently, it was shown that altering the design of the targeting vector so that the homology arms are flanked by sgRNA+PAM sequences improved the efficiency of HDR in both HEK293T cells and iPSCs. This efficiency was also further improved by treating cells with small molecules shown to regulate the cell cycle.[91]

In summary, we have used the CRISPR/Cas9 system to perform various genome edits to the *UBE3A* genomic sequence, generating isogenic stem cell lines. These isogenic lines have been used to establish robust cellular models for studying AS and *UBE3A*, and for establishing and validating several AS *in vitro* phenotypes.

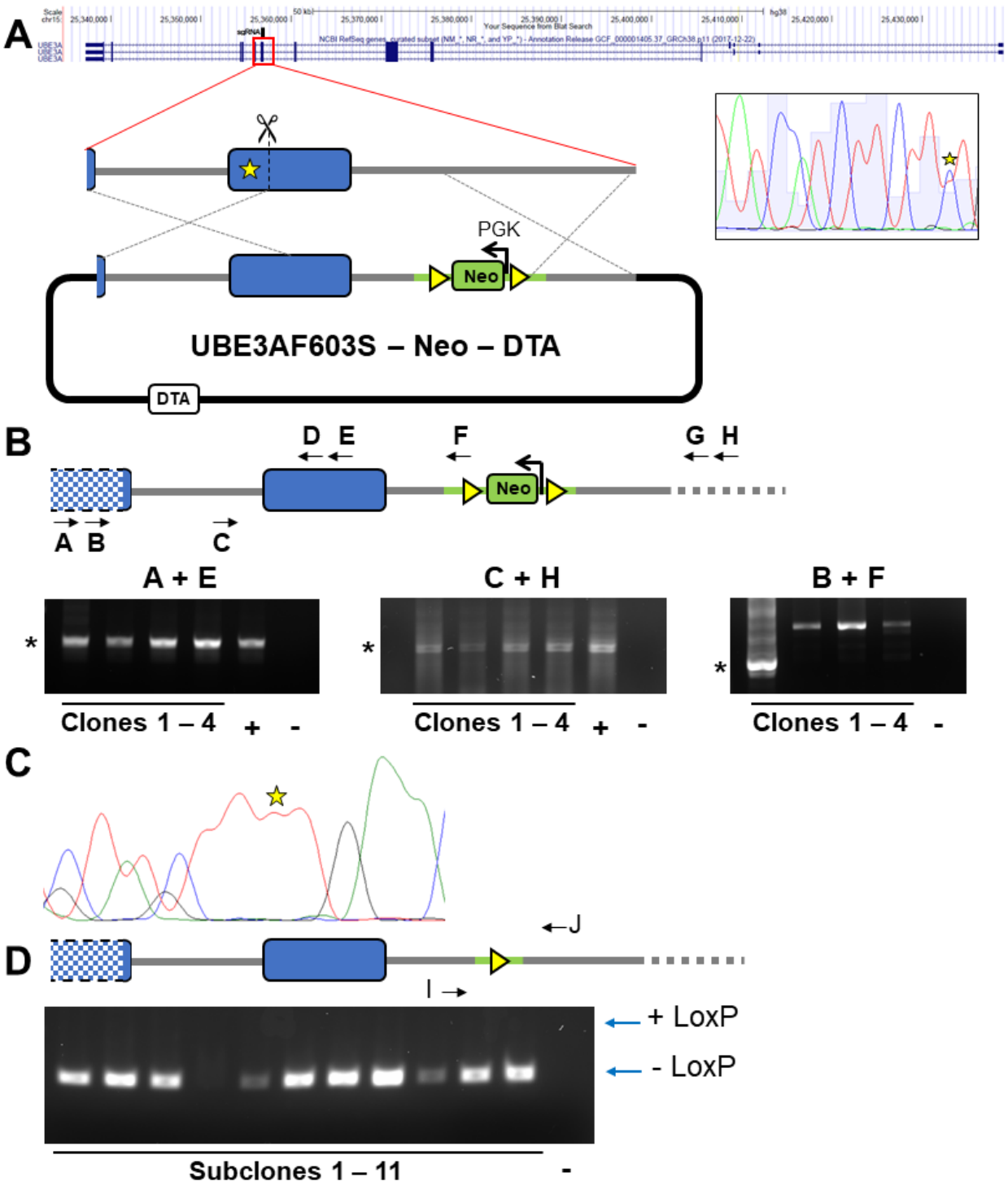


Figure 1. Correction of AS point mutation using CRISPR and targeting vector
A Top: location of proposed genome editing in UBE3A gene. sgRNA = location of sgRNA used to correct mutation. Bottom: schematic illustrating proposed genome editing using a targeting vector as the template for HDR. Inset: Sanger sequencing of iPSCs containing point mutation. **B** Top: schematic illustrating correctly targeted allele and PCR primers used for screening clones… (continued on next page)

Bottom: Agarose gel images showing PCR products in clones screened. **C** Sanger sequencing of clone 1 showing correction of point mutation **D** Top: schematic illustrating correctly targeted clone following Cre-mediated recombination and removal of Neo cassette; Bottom: Agarose gel image showing PCR products in clones screening following nucleofection with Cre-expressing plasmid.

Scissors = Cas9 cut site; yellow star = location of point mutation; blue box = exon; grey line = intron; solid line/box = region used for homology in targeting vector; dashed line = genomic sequence outside of homology region; yellow triangle = loxP; Neo = neomycin resistance cassette; PGK = promoter; DTA = diptheria toxin A; arrow = primer location; + = positive control; - = no template control; * = expected product size

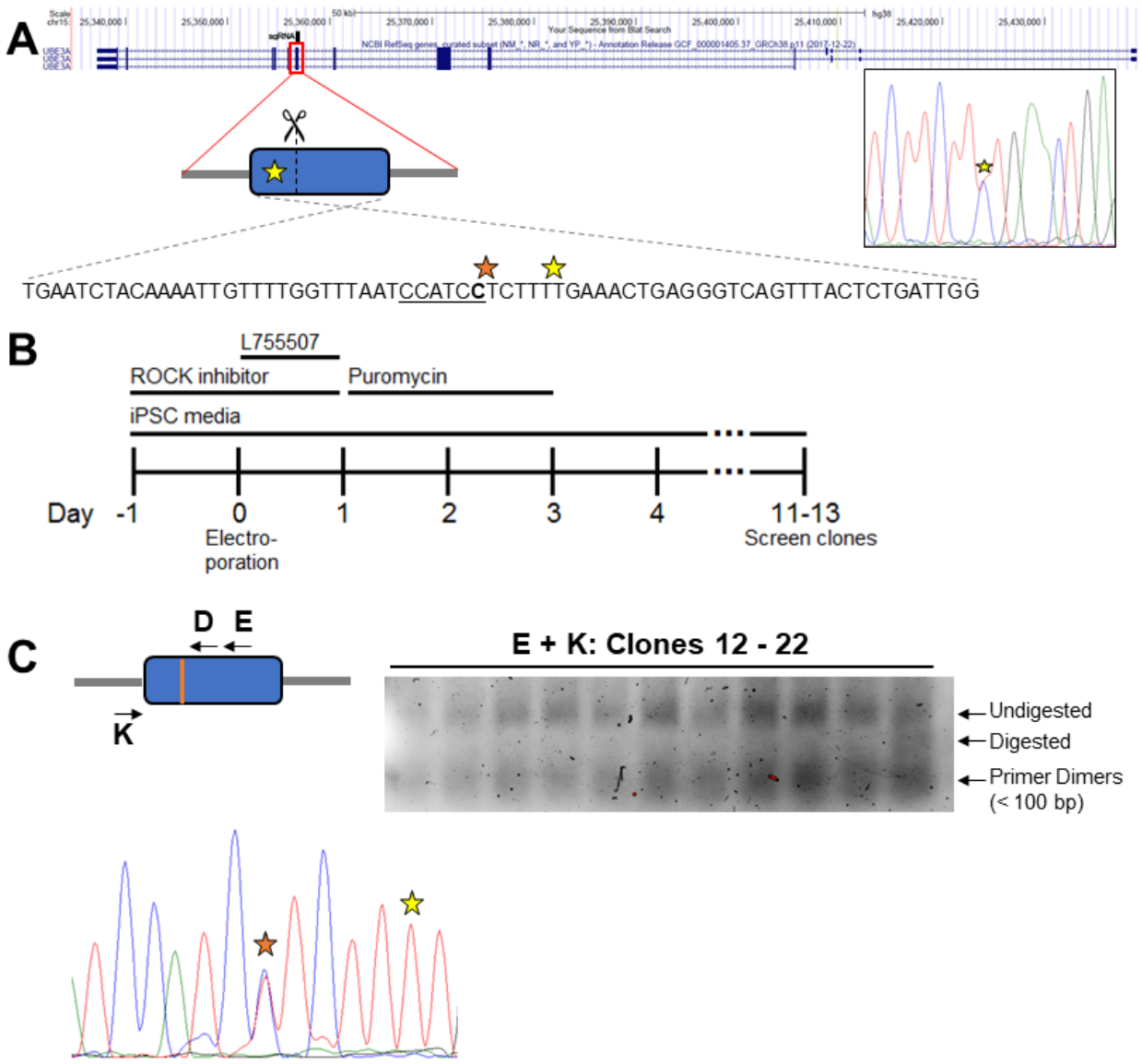
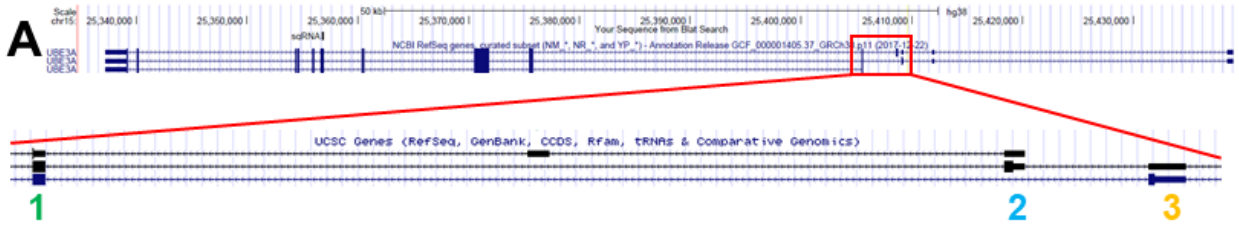


Figure 2. Correction of AS point mutation using CRISPR and ssODN
A Schematic showing proposed genome editing in UBE3A gene and sequence of ssODN used as template for HDR. sgRNA = location of sgRNA used to correct mutation. Inset: Sanger sequencing of AS point mutation in patient-derived iPSCs **B** Schematic showing timeline for genome editing experiments. **C** Top left: schematic showing locations of primers in targeted clone. Orange line indicates location of silent mutation, arrows indicate primers. Top right: Agarose gel image showing digested PCR products from clone screen following FokI digest. Bottom: Sanger sequencing of clone 22 showing correction of point mutation.
Yellow star = point mutation location; orange star = location of silent mutation

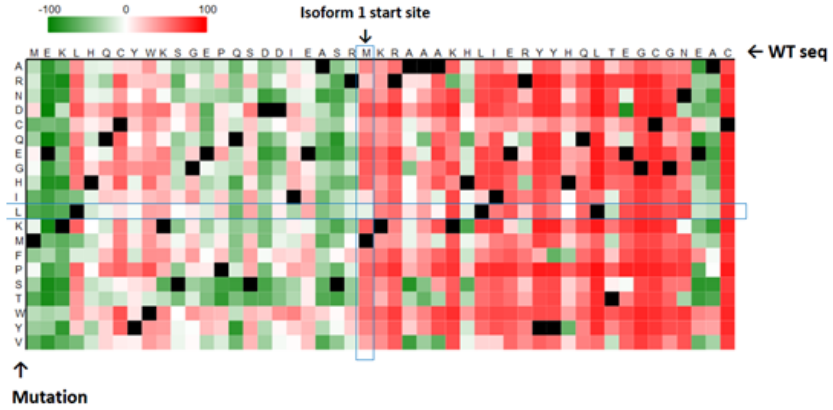


Isoform 1: MKRAAAKHLIERYYHQLT….

Isoform 2: MEKLHQCYWK**SGEPQSDDIEASR**MKRAAAKHLIERYYHQLT… (+23 AA)

Isoform 3: MATACKR**SGEPQSDDIEASR**MKRAAAKHLIERYYHQLT… (+20 AA)

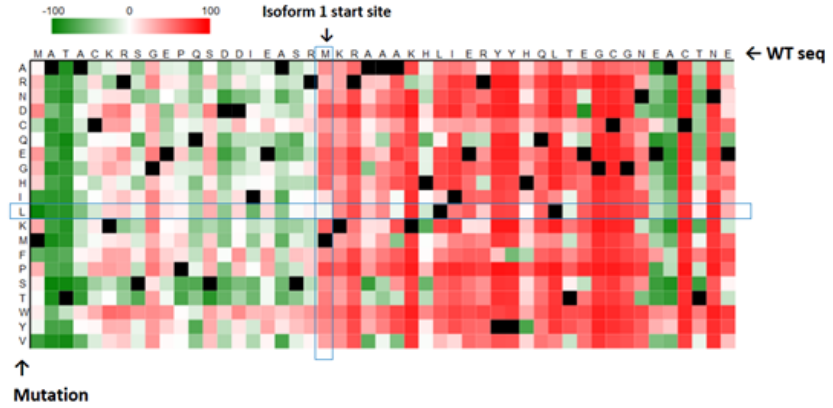
B Isoform 2 – Predict Protein



Isoform 2 – Predict SNP

RESULTS			neutral	deleterious	XX % confidence		Expand all annotations	
Annotation	Mutation	PredictSNP	MAPP	PhD-SNP	PolyPhen-1	PolyPhen-2	SIFT	SNAP
	M24L	83 %	77 %	66 %	67 %	70 %	71 %	55 %

C Isoform 3 – Predict Protein



Isoform 3 – Predict SNP

RESULTS			neutral	deleterious	XX % confidence		Expand all annotations	
Annotation	Mutation	PredictSNP	MAPP	PhD-SNP	PolyPhen-1	PolyPhen-2	SIFT	SNAP
	M21L	83 %	77 %	66 %	67 %	74 %	71 %	55 %

Figure 3. Predicted effects of proposed mutation at isoform 1 translation start site on protein structure of isoforms 2 and 3 … (continued on next page)

A Top: Location of isoform translational start sites. Bottom: Amino acid sequence of isoform N-termini. Red letters indicate common amino acids between the three isoforms. M indicates methionine used as translational start site. **B & C** Effect of proposed amino acid change on isoform 2 (**B**) and isoform 3 (**C**) protein. Top: results from Predict Protein software showing effects of every possible amino acid substitution of the methionine. Bottom: results from Predict SNP software showing predicted effects of L to M substitution from 7 different structure prediction algorithms.

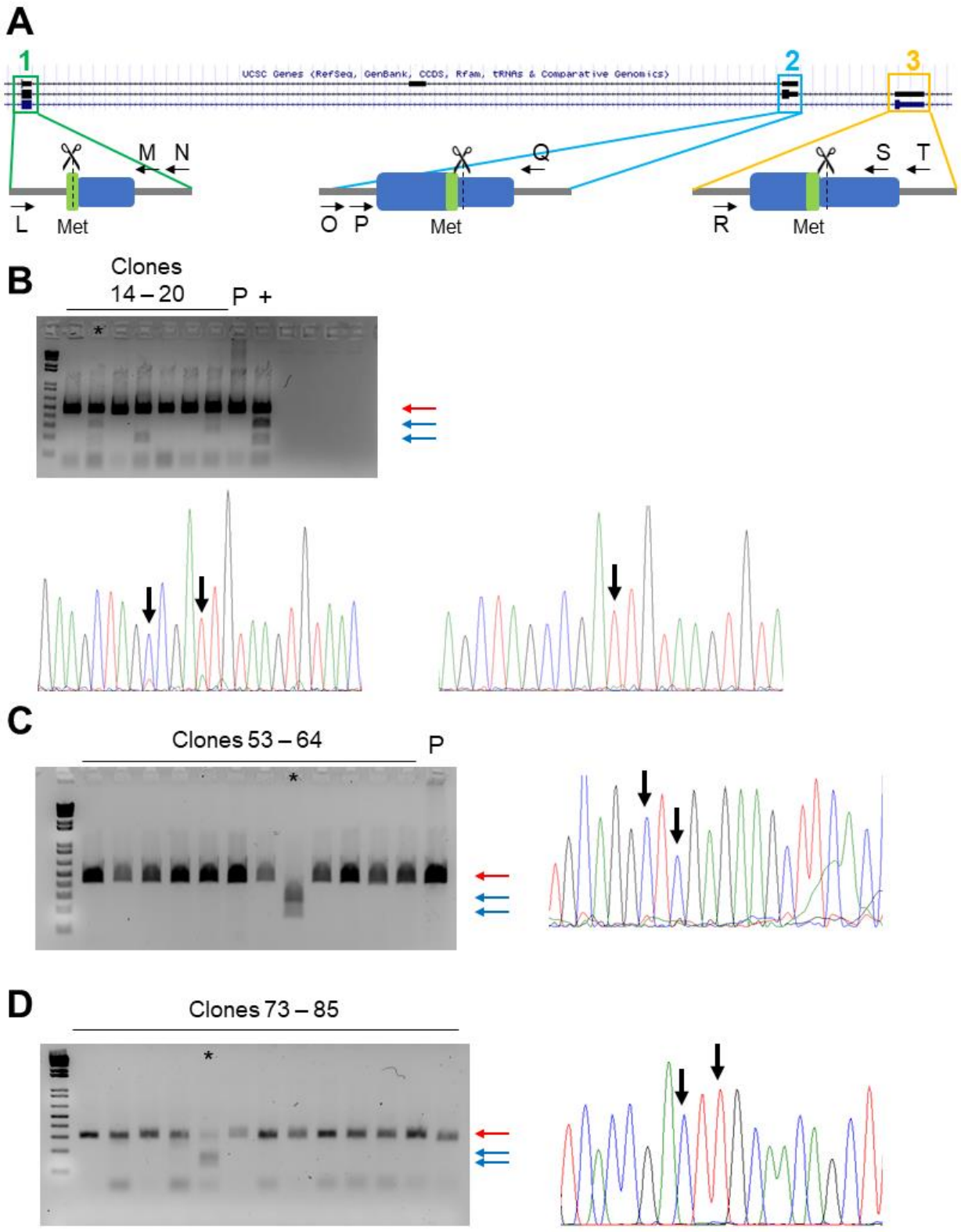


Figure 4. Generation of isoform-null hESC lines

A Schematic illustrating proposed genome edits at each isoform translational start site. Grey line = intron; blue box = exon; … (continued on next page)

green box = methionine used as start site; scissors = Cas9 cut site; arrows = primers used for screening **B** Generation of isoform 1-null line. Top: agarose gel image showing TaqI restriction digest of PCR products from clones. Bottom left: Sanger sequencing of clone 15 indicated that it was a mixed clone (arrows). Bottom left: Sanger sequencing of clone 15 after routine passage showing absence of WT sequence and presence of homozygous start site mutation (A to T). **C** Generation of isoform 2-null line. Left: agarose gel image showing HindIII restriction digest of PCR products from clones. Right: Sanger sequencing of clone 60 showing homozygous silent mutation and translational start site mutation. **D** Generation of isoform 3-null line. Left: agarose gel image showing HPYCHV restriction digest of PCR products from clones. Right: Sanger sequencing of clone 77 showing homozygous silent mutation and translational start site mutation.

* = clone used for cell line and clone shown in Sanger sequence trace; P = parent line (H9 hESCs); + = positive control for restriction digest; red arrow = undigested PCR product; blue arrows = digested PCR product; black arrows = edited nucleotides

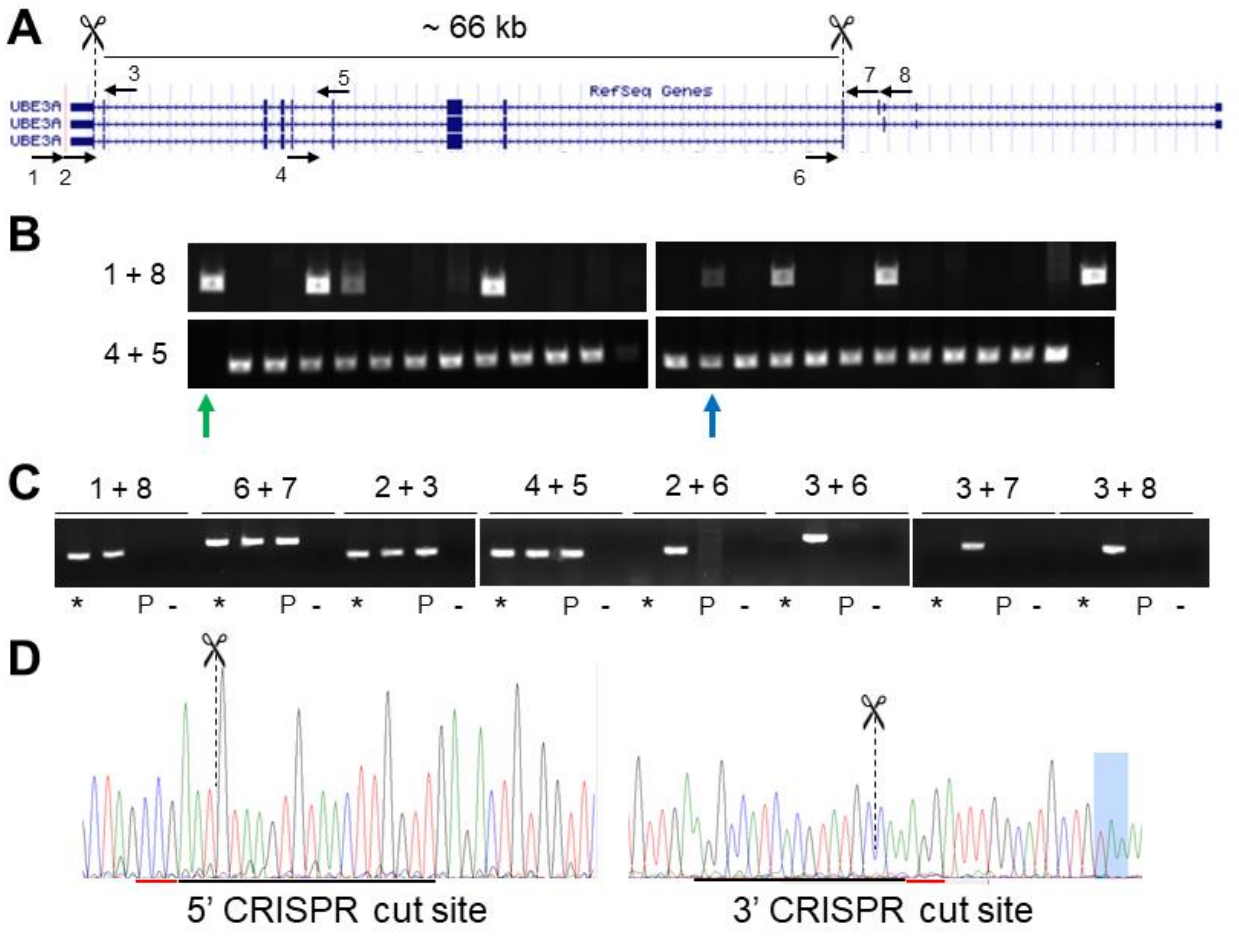


Figure 5. Generation of UBE3A deletion in hESCs
A Schematic illustrating proposed genome editing. Scissors = Cas9 cut sites. Arrows = locations of primers used for screening clones **B** Agarose gel images showing PCR products from indicated primer combinations. Top: primers spanning the deletion region: positive product indicates deletion occurred on at least one allele Bottom: primers located within deleted region: positive product indicates at least one copy of UBE3A is intact. Green arrow shows a clone with a homozygous deletion, blue arrow shows a clone with a heterozygous deletion **C** Characterization of heterozygous clone via various combinations of primers (explained in **Table 3**). * = heterozygous clone; P = parent line; - = no template control **D** Sanger sequencing of heterozygous clones showing no mutations at the CRISPR cut sites on the intact allele. Scissors indicate Cas9 cutting location. Red line indicates PAM (NGG). Black line indicates sgRNA binding location. Left: cut site at 5' end of *UBE3A* gene. Right: cut site at 3' end of *UBE3A* gene. The stop codon is highlighted in blue.

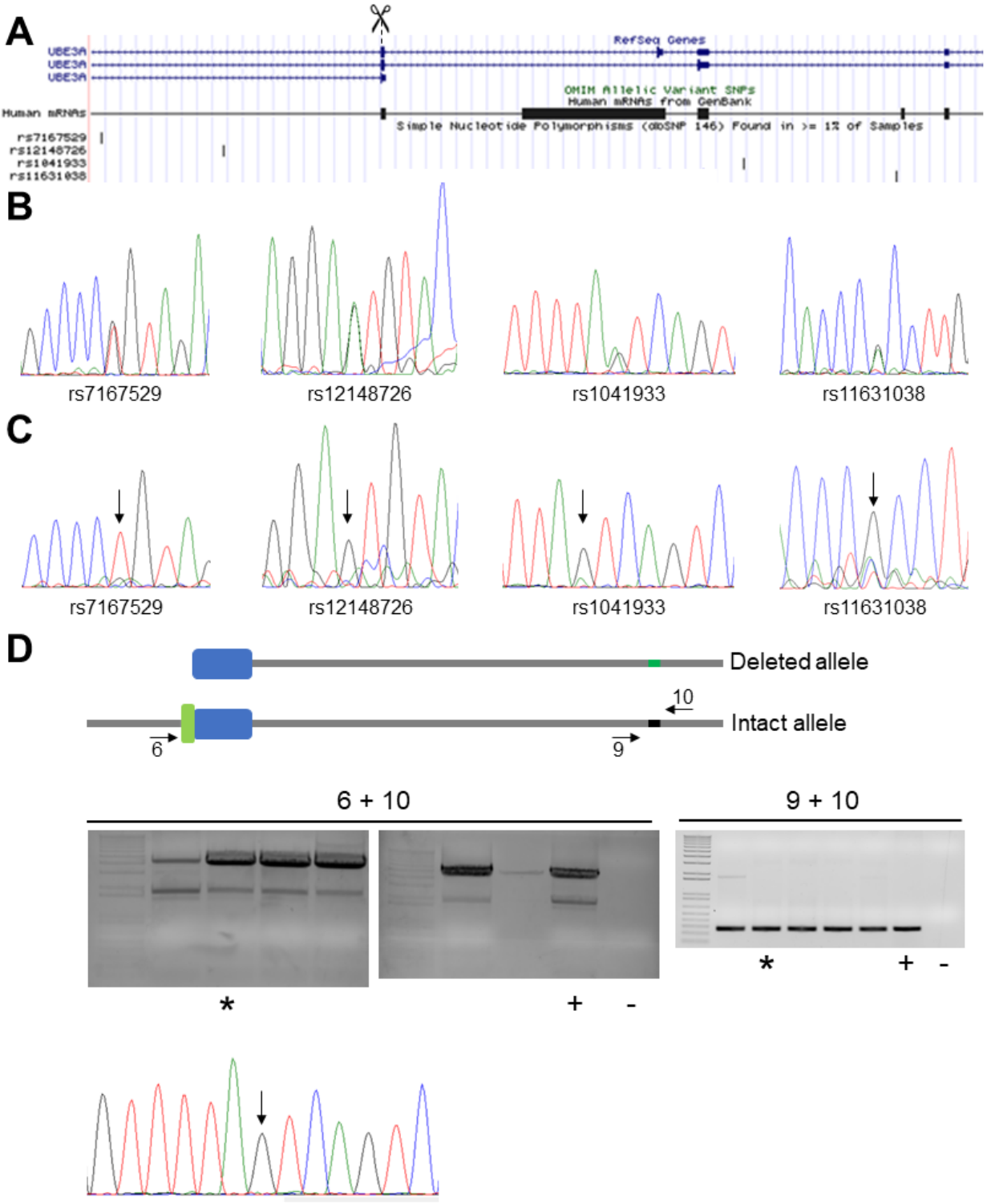


Figure 6. Determining the identities of the deleted and intact alleles
A Location of SNPs relative to 5' CRISPR cut site. **B** Sanger sequencing of genomic DNA from H9 hESC showing presence of 4 SNPs. **C** Sanger sequencing of *UBE3A-ATS* transcript cDNA showing presence of 4 SNPs. Arrow indicates location of SNP. … (continued on next page)

D Top: schematic illustrating sequencing of deleted and intact alleles in heterozygous UBE3A deletion clone. Horizontal arrows indicate primer locations. Green and black segment indicates location of SNP. Middle: agarose gel images showing PCR amplification of non-deleted allele (left) followed by amplification around the SNP of interest (right). * = PCR product from heterozygous deletion clone; + = positive control; - = no template control; Bottom: Sanger sequencing showing identity of the SNP (vertical arrow).

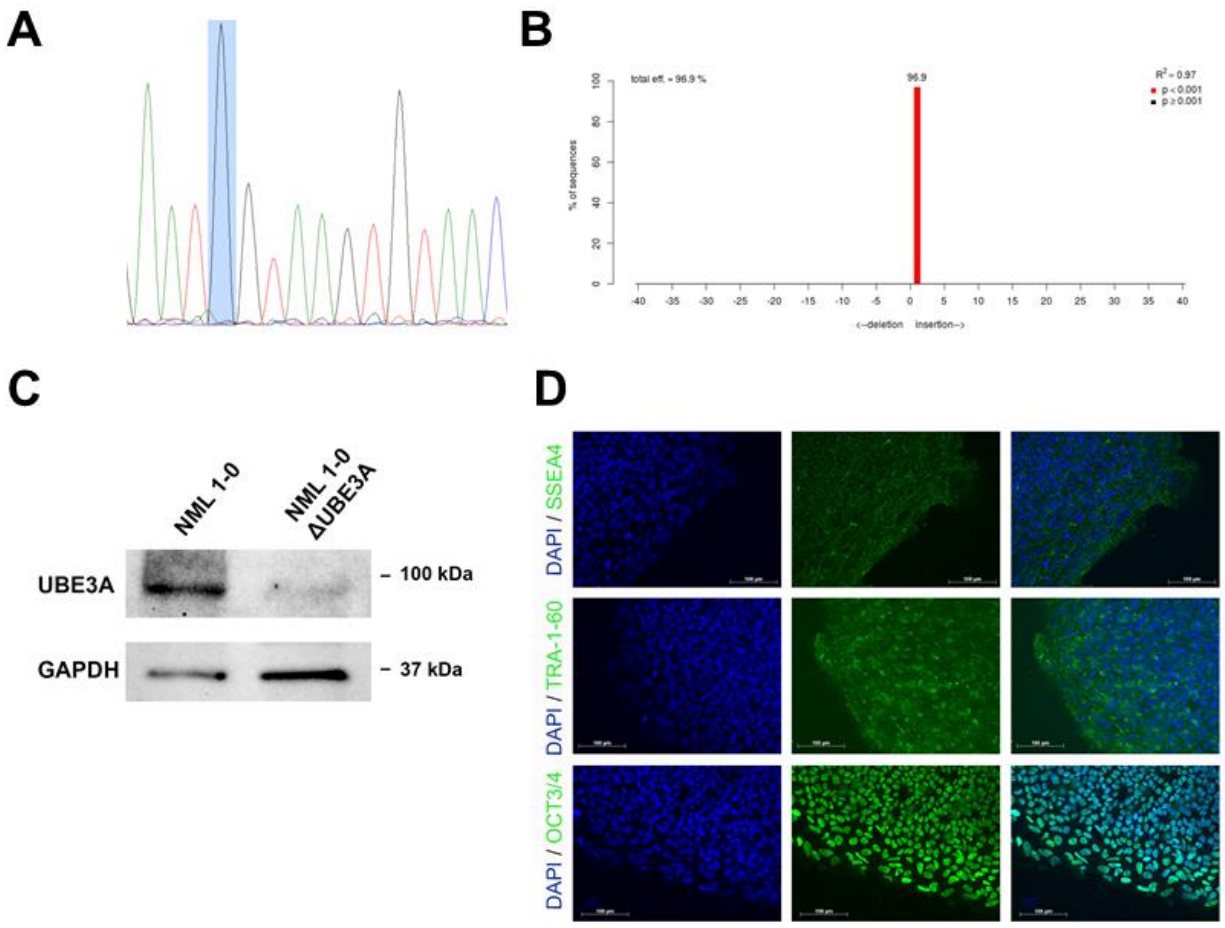


Figure 7. UBE3A knockout in normal iPSCs
A Sanger sequencing showing insertion of a G at CRISPR cut site (blue highlighted nucleotide) **B** TIDE analysis of clone after singlization showing no wildtype sequence and insertion of 1 bp **C** Western blot comparing parent iPSC line (NML1-0) to UBE3A KO iPSC line (NML1-0 ΔUBE3A) **D** Editing of normal iPSCs did not affect pluripotency of iPSCs, as indicated by expression of multiple pluripotency markers.

Table 1. sgRNAs used to generate stem cell lines		
Cell Line/Project	sgRNA sequence + **PAM**	Vector
Correct mutation in Rx35i-7 iPSCs	CCCTCAGTTTCAGAAGAAGA**TGG**	pX330
Correct mutation in Rx35i-7 iPSCs	CCCTCAGTTTCAGAAGAAGA**TGG**	plentiCRISPRv2
Isoform 1 start site mut. (H9); UBE3A KO in MCH2-10; UBE3A deletion (H9)	**CCG**AATGTAAGTGTAACTTGGTT	pX459v2
Isoform 2 start site mut. (H9)	AAAAGGAGTGGCTTGCAGGA**TGG**	pX459v2
Isoform 3 start site mut. (H9)	ATCACCCTGATGTCACCGAA**TGG**	pX459v2
3' CRISPR for UBE3A deletion (H9)	AGGCCATCACGTATGCCAA**AGG**	pX330

Table 2. ssODNs used to generate stem cell lines	
Cell Line/Project	ssODN sequence
Correct mutation in Rx35i-7 iPSCs	TGAATCTACAAAATTGTTTTGGTTTAATCCATCCTCTTTTGAAACT GAGGGTCAGTTTACTCTGATTGGC
Isoform 1 start site mutation (H9)	AGAACCTCAGTCTGACGACATTGAAGCTAGTCGATTGTAA GTGTAACTTGGTTGAGACTGTGGTTCTTAT
Isoform 2 start site mutation (H9)	ATTCAAATGGTGGCTCACTTCCAATAACACTGGTGAAGCTTCTCGA GCCTGCAAGCCACTCCTTTTACCTCCACTGTAACTCTCTAGGAGAG
Isoform 3 start site mutation (H9)	TGTAAAATAATTCAAAATTACCTTTTACAAGCTGTTGCAAGTCGGTG ACATCAGGGTGATCACAGCTTTGAGTCACTGATTAAAAA

Table 3. Primer combinations used to screen UBE3A deletion clones	
Primer Set	PCR Result Interpretation
1 + 8	Presence of band indicates deletion of at least one allele
6 + 7	Use to sequence 5' cut site of intact allele to check for indels
2 + 3	Use to sequence 3' cut site of intact allele to check for indels
3 + 7	Presence of band indicates inversion of intact allele (two forward primers)
3 + 8	Presence of band indicates inversion of intact allele (two forward primers)
4 + 5	Presence of band indicates presence of at least one intact allele / absence of band indicates a homozygous deletion
2 + 6	Presence of band indicates inversion of intact allele (two reverse primers)
3 + 6	Presence of band indicates a duplication of the intact allele

Table 4. Primer Sequences Used for Genome Editing		
Primer Name	Letter/Number in Fig.	Sequence
EndogR2	A	TTGGAATTTTATCACCATTTTCCA
EndogR1	B	GATCATCATGTCATCTTCCACATT
OrigR	C	TTTTCCCCATTAGCTTCCT
OrigSeqF	D	AGGTATGTTCACATACGATG
OrigF	E	TTCATCAGTCCTTAATAAAATACAAAA
ExogF2	F	ACGAAGTTATTAGGTCCCTCG
EndogF2	G	CAACTACTCCGGAGGCTGAG
EndogF1	H	GGATTACAGGTGTGAACTACCACA
EnNeoR	I / 4	GCACTTGAGAAAACAATGTCCA
EnNeoF	J / 5	TGAAACACTTTGGAAATGTAGCC
NewR	K	TTATTGTAAATAGCCAGACCCAGT
Iso1R	L / 6	CTGCTACCAGGGAAGCAAAA
Iso1SeqF	M	TTCTTTCATGTTGACATCTTTAATTTT
Iso1F	N / 7	GCTTATAATGGCTTGTCTGTTGG
Iso2R	O	TGAAACAATAACCAAATAACATTGG
Iso2SeqR	P	TCTTGATTTGAATCGCAGAAAA
Iso2F	Q	TCAGTAGCCACTATCAAAGACCT
Iso3R	R	TTTTTGAACAATGAATTGGGTTT
Iso3SeqF	S	AGCCTACGCTCAGATCAAGG
Iso3F	T	TTTTTGAACAATGAATTGGGTTT
UBE3A_del_F1	1	AGTCCAACCCTTAAAATAAATGTG
UBE3A 3’ R	2	TGGGACACTATCACCACCAA
UBE3A_intron12_F1	3	GGCAACTTGGTAGTTACACAACA
UBE3A_del_R1	8	CCCACATGTCCCCAATAAAG
SNP-933 R	9	AAAATTAAGCAGCCTCCGAGT
SNP-933 F	10; Also used for ssRT	TTAAGCAGTTGCCCTCCTTG
Primers also used but not shown in any figures:		
UBE3A_intron12_F2	seq 3’ cut site	CCCATGACTTACAGTTTTCCTG
SNP-933 seq	seq SNP	TGACTGAATTTGTCCGTATTTGA
SNP-1038 F	Also used for ssRT	GCAGTGGCACCTTCTTGACT
SNP-1038 R		GAGCTGTCATGGATAAACAAGC
SNP-1038 seq	seq SNP	CCTCTCTTCCAGTCCCCTCT
SNP-726 F	Also used for ssRT	TTCACAGCCTCTGACAACCA
SNP-726 R		CACAATATGTGATGGCAAGGA
SNP-726 seq	seq SNP	GCTCTGGAAGAGTATGGCAGT
SNP-529 F	Also used for ssRT	TGCAGAACAAGGGTCCAAAC
SNP-529 R		GGTGTCACGTTTCCCTGACT
SNP-529 seq	seq SNP	ACACTGGTCCTTGCAGATCC

Chapter 3

Establishing a quantitative, molecular phenotype for Angelman Syndrome pluripotent stem cell-derived neurons

Carissa L. Sirois[1], Dea Gorka[1], James J. Fink[2], Noelle D. Germain[1], Fabian Offensperger[3], Martin Scheffner[3], Eric S. Levine[2], Stormy J. Chamberlain[1]

[1] Department of Genetics and Genome Sciences, UConn Health, Farmington, CT, USA
[2] Department of Neuroscience, UConn Health, Farmington, CT, USA
[3] Department of Chemical Biology, University of Konstanz, Konstanz, Germany

This chapter is a manuscript being prepared for submission.

My contributions: Work in this chapter was performed by the author aside from karyotyping of stem cells, whole-cell patch clamp recordings, *in vitro* ubiquitination assays, and treatment of normal neurons with antisense oligonucleotides

Abstract

Angelman Syndrome (AS) is a neurodevelopmental disorder occurring approximately once in every 15,000 live births, characterized by severe seizures, absent speech, motor dysfunction, profound intellectual disability, and happy demeanor. Loss of expression of the maternal copy of *UBE3A*, a gene regulated by tissue-specific genomic imprinting, causes AS. *UBE3A* encodes an E3 ubiquitin ligase that may also act as a transcriptional co-activator. Although there is currently no cure, multiple therapies for AS, including gene therapy, are currently being explored. It will therefore be necessary to test the efficacy of these therapies in human AS neurons. Here, we have used isogenic AS and control stem cell-derived neurons to establish a molecular transcriptome phenotype for AS. We have generated two AS/control isogenic stem cell line pairs. First, we have derived induced pluripotent stem cells (iPSCs) from an AS patient with a missense mutation in *UBE3A* (AS Pt Mut). This mutation does not affect UBE3A RNA or protein levels but causes a reduction in the protein's ubiquitin ligase activity, as demonstrated by *in vitro* ubiquitination assays. Using CRISPR/Cas9-mediated genome editing, we have corrected this point mutation in the AS Pt Mut iPSCs, generating an isogenic control iPSC line (Corrected). Second, we have used CRISPR/Cas9 to delete the maternal copy of UBE3A in H9 human embryonic stem cells (hESCs). Both isogenic pairs of stem cells were then successfully differentiated into forebrain neurons. Whole-cell patch-clamp recordings were performed on 12-week neurons to determine the extent to which the AS stem cell-derived neurons recapitulate the electrophysiological phenotypes seen in other non-isogenic AS iPSC-derived neuron lines. Stand-specific mRNAseq was then performed on all 4 neuron lines to find a list of common genes differentially expressed in AS neurons ($n = 855$, $p < 0.01$). Lastly, we tested the ability to rescue electrical and molecular phenotypes by treating AS Pt Mut and Corrected neural progenitors with lentivirus encoding E6 oncoprotein, which restores ubiquitin ligase function to the mutant version of UBE3A in *in vitro* ubiquitination assays. Expression of E6 in point mutation neurons was able to rescue both the resting membrane potential electrical phenotype as well as

restore expression of multiple genes identified by mRNAseq. These results demonstrate that the transcriptome phenotype is a useful tool assaying phenotypic restoration in stem-cell derived neurons, and also suggest that allosteric activation of UBE3A may be a useful mechanism to explore for potential AS therapies.

Introduction

Chromosome 15q11-q13 is regulated by genomic imprinting, an epigenetic phenomenon that results in the differential expression of genes in a parent of origin-specific manner.[1] The *UBE3A* gene found in this region undergoes tissue-specific imprinting: it is expressed biallelically in non-neuronal cells, but is silenced on the paternal allele in neurons due to the expression of a long-noncoding RNA, the *UBE3A-antisense (UBE3A-ATS)*.[11] Subsequently, mutations or deletions affecting the maternal copy of *UBE3A* result in a loss of *UBE3A* in neurons, which causes Angelman Syndrome (AS).[21] AS is a rare neurodevelopmental disorder that occurs in 1 in 15,000 live births, and is characterized by severe seizures, ataxia, happy affect, absent speech, and intellectual disability.[20] Although there is no cure for AS, several promising approaches to restore UBE3A expression are currently being explored as potential AS therapies. Unsilencing the paternal copy of *UBE3A,* using either topoisomerase inhibitors or antisense oligonucleotides, has been demonstrated in *in vitro* and *in vivo* studies, and was shown restore some phenotypes in adult AS mice.[65,70] Viral vector-based gene therapy to replace the *UBE3A* gene product is also being explored.[71,72] While demonstrating the safety and efficacy of these approaches in mice is important, it is also necessary to test these potential therapies in human cells, specifically human neurons, not only to assess their toxicity but to assess the relative effectiveness of therapies that are specific to the human genomic sequence. Human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs), which can be differentiated into neurons, provide us with the unique opportunity to test these potential therapies on human neurons *in vitro*. In order to do so, however, it is necessary to first establish robust *in vitro* AS phenotypes that can be assayed following treatment of neurons with potential therapeutics. Previously, we showed that AS iPSC-derived neurons exhibit a phenotype of impaired electrophysiological maturation.[76] Here, we have used CRISPR/Cas9 to generate two pairs of isogenic AS and control human pluripotent stem cells. We established a quantitative molecular phenotype for AS neurons by performing mRNAseq on two pairs of isogenic AS and

control stem cell-derived neurons. Using conservative criteria, we have identified a list of genes that are consistently differentially expressed in AS that can serve as a molecular phenotype for human stem cell-derived AS neurons. We validated the use of these genes as a molecular phenotype by assaying their expression following restoration of UBE3A function in neurons differentiated from AS iPSCs. We propose that this list of genes can be used as a tool to assess the efficacy of potential AS treatments when testing them on cultured human AS neurons. These genes can be assayed in a relatively high-throughput manner, making the transcriptome an ideal phenotype for screening potential AS therapeutics.

Methods

Patient-Specific Fibroblasts & Generation of iPSCs

Fibroblasts were obtained from a skin punch biopsy following appropriate informed consent from a 7-year old male with Angelman syndrome. This boy carried a C to T transition resulting in a phenylalanine to serine missense mutation in the maternally-inherited copy of his UBE3A gene. The parent of origin of the mutation was known because the mother and maternal grandfather of the boy both carried the same mutation. iPSCs were derived by the UConn-Wesleyan Stem Cell Core (Farmington, CT) from patient fibroblasts using the human polycistronic STEMCCA lentiviral vector encoding *OCT4, KLF4, SOX2,* and *CMYC,*[92] as previously described.[93]

iPSC/hESC culture and neuronal differentiation

iPSCs and hESCs were cultured on irradiated mouse embryonic fibroblasts (irrMEFs) and fed daily with SC media consisting of DMEM/F12 containing knockout serum replacement, L-glutamine + β-mercaptoethanol, non-essential amino acids (NEAA), and basic fibroblast growth factor. SCs were manually passaged every 5-7 days and cultured at 37°C in a humid incubator at 5% CO_2.[78]

SCs were differentiated into neurons using a modified version of the monolayer differentiation protocol.[79,94] Two days after passaging, SCs were cultured in N2B27 medium (Neurobasal medium, 1% N2, 2% B27, 2 mM L-glutamine, 0.5% penicillin/streptomycin, 1% insulin-transferrin-selenium). N2B27 medium was supplemented with fresh Noggin (500 ng/mL) for the first 10 days of differentiation. Approximately 2 weeks after beginning differentiation, neural rosettes were passaged onto poly-ornathine and laminin-coated plates using the StemPro EZ Passage tool (Thermo Fisher). Neural progenitors were replated at a high density around 3 weeks of differentiation, switched to neural differentiation medium (NDM) around 4 weeks of differentiation, then plated for terminal differentiation around 5 weeks. NDM consisted of neurobasal medium, 1% B27, 2 mM L-glutamine, 0.5% pen-strep, non-essential amino acids, 1

μM ascorbic acid, 200 μM cyclic AMP, 10 ng/mL brain-derived neurotrophic factor, and 10 ng/mL glial-derived neurotrophic factor. Neurons were maintained in culture until 10 to 12 weeks of differentiation.

CRISPR/Cas9-mediated genome editing

sgRNAs were designed using MIT's CRISPR Design Tool (http://crispr.mit.edu) and cloned into either plentiCRISPRv2 (Addgene 52961) or pX459v2.0 (Addgene 62988) vector, as previously described.[95,96] The sgRNA used to correct AS Pt Mut iPSCs and the 3' gRNA used to delete UBE3A in hESCs were designed by the hESC/iPSC Targeting Core at the University of Connecticut Health Center. Prior to electroporation or nucleofection, SCs were treated with ROCK inhibitor (Y-27632; Selleck Chemicals) for 24 hours. SCs were then singlized using Accutase (Millipore) and electroporated using the Gene Pulser X Cell (BioRad) or nucleofected using the Amaxa 4D Nucleofector (Lonza). For generation of the Corrected line, iPSCs were electroporated in PBS with 10 μg of CRISPR plasmid and 8 μl of single-stranded oligonucleotide (ssODN) template. For generation of the ΔUBE3A$^{m-/p+}$ line, hESCs were nucleofected with 2.5 μg of each CRISPR using the P3 Primary Cell Kit L (Lonza). SCs were then plated onto puromycin-resistant (DR4) irrMEFs at low density, supplemented with ROCK inhibitor and L755507 (5 μM, Xcessbio), which has been shown to improve efficiency of homology directed repair.[84] 24 hours after plating, cells underwent selection for 48 hours with puromycin (0.5 – 1 μg/ ml). Puromycin resistant colonies were screened by conventional PCR (UBE3A deletion) or conventional PCR followed by restriction digest (AS Pt Mut correction) approximately 11-14 days after plating. Putative clones were plated onto regular irrMEFs and successful genome editing was confirmed by Sanger sequencing.

Conventional PCR and Stand-specific RT-PCR

Conventional PCR was performed on genomic DNA using Herculase II Fusion DNA Polymerase (Agilent). Strand-specific RT using primers specific to the UBE3A-ATS transcript was performed on RNA isolated from 10-week-old H9 hESC-derived neurons using Superscript III Reverse Transcriptase (Thermo Fisher Scientific). RT-PCR was performed on the strand-specific cDNA libraries using Advantage 2 Polymerase (Clontech).[97]

Karyotype and pluripotency analysis

Normal karyotype of AS Pt Mut, Corrected, & ΔUBE3A$^{m-/p+}$ SCs was confirmed by the Genetics and Genomic Division of the University of Connecticut-Wesleyan University Stem Cell Core (**Supplemental Figures S1-S2**). A minimum of five G-banded metaphase stem cells were examined from each cell line.

Electrophysiology

Neurons were plated onto glass coverslips between 5 and 6 weeks of differentiation. Whole-cell voltage and current clamp recordings of 11- to 13-week-old stem cell derived neurons were performed as previously described.[76] Briefly, recordings were performed on individual coverslips in a recording chamber fixed to the stage of the microscope that was continuously perfused with oxygenated artificial cerebrospinal fluid. Neurons were identified based on their morphology when viewed on infrared differential interference contrast (DIC) video microscopy. Patch pipettes (3 to 5 Ω) were pulled from borosilicate glass capillaries and filled with internal solution containing KCl, K-gluconate, HEPES, phosphocreatine, EGTA, $CaCl_2$, Na_2-ATP, and Na-GTP. Input resistance (R_i) was noted throughout the recording period. The following criteria were used to exclude neurons from analysis: series resistance (R_s) > 50 MΩ at time of break-in, R_i changed by > 15% during recording, or R_i < 100 MΩ during recording. Resting membrane potential (RMP) was noted at break-in by injection of 0 current. RMP values were corrected for liquid junction potential offline. To elicit AP firing, cells were held in current clamp mode at

approximately -70 mV, followed by application of 500 ms current steps from -20 to +40 pA in intervals of 5 pA. To elicit inward and outward currents, cells were held in voltage-clamp mode at -70 mV, followed by application of 300 ms steps from -100 to +40 mV in 10 mV increments. Finally, spontaneous excitatory synaptic activity was monitored while in voltage-clamp mode by holding the cell at -70 mV. Offline analysis of data was done using Clampfit software. For E6 experiments, recordings were performed as described above with the following changes: after identifying neurons by morphology in infrared DIC, cells were then viewed in the GFP channel to find neurons that expressed GFP. GFP+ neurons were confirmed using TurboSM (RedShirtImaging) at the following settings: exposure time 100 sec, number of frames 1000, camera gain Low. Following confirmation of GFP expression, recordings were then performed while viewing the cell in infrared DIC.

RNA sequencing

Total RNA was isolated from 6 biological replicates of 12-week-old neurons differentiated from AS Pt Mut iPSCs, Corrected iPSCs, H9 hESCs, and H9ΔUBE3A$^{m-/p+}$ hESCs using RNA Bee (Amsbio). ASO-treated H9 neurons were collected at the end of treatment (approximately 10.5 weeks of differentiation). mRNAseq libraries were prepared using the TruSeq Stranded mRNA Library Prep Kit (Illumina) per the manufacturer's protocol. 500 ng of RNA was used as input for each sample. Libraries were sequenced using the NextSeq500 sequencer (Illumina). Fastq files were run through fastqc and visualized using MultiQC, which revealed adapter contamination in a handful of samples. All samples were run through Trimmomatic before proceeding with analysis. Reads were aligned to the human genome (GRCh37/hg19) and gene counts extracted using STAR. For AS Pt Mut, Corrected, H9, & ΔUBE3A$^{m-/p+}$ neurons, at least 30 million reads were generated for each sample. For the ASO-treated H9 neurons, at least 18.4 million reads were generated for each sample. With the exception of one AS Pt Mut replicate (AS1; 74.82%), all samples had at least 84.24% uniquely mapped reads. Zero count genes and the bottom 10%

of genes with counts were removed, then counts normalized and transformed. Data was subjected to surrogate variable analysis[98] then differential expression analyzed using DESeq2. For each analysis, an adjusted p-value (padj) of $p<0.01$ was used as a cut off for differential gene expression. Gene ontology analysis was performed using DAVID. Trimmomatic percentages, and alignment percentages can be found in **Supplemental Table S3**. Software versions can be found in **Supplemental Table S4**.

Acute UBE3A knockdown experiment

Ten-week-old neurons differentiated from H9 hESCs were treated for 72 hours with 10 μM of antisense oligonucleotide (ASO) against *UBE3A* or scramble control ASO. After treatment, neurons were harvested for RNA isolation. ASOs were provided by Ionis pharmaceuticals and have been described previously.[76]

Site-directed mutagenesis of UBE3A cDNA for *in vitro* ubiquitination assay

The AS-causing point mutation seen in the AS Pt Mut iPSCs was generated in UBE3A cDNA (c7-3)[99] that had been cloned into pBluescript (SK-) plasmid. To generate a template containing the point mutation and restriction digest sites for ligation, sequential PCR reactions were performed off of the plasmid using restriction site-containing or mutation-containing primers. The PCR product and plasmid were digested with restriction enzymes, gel purified, then ligated. The resulting plasmid was transformed in DH10B chemically competent cells. Bacterial clones were screened by Sanger sequencing to confirm presence of the point mutation in the cDNA.

In vitro ubiquitination assay

Wildtype and mutated UBE3A cDNA were in vitro translated as previously described.[41] To examine the auto-ubiquitination of UBE3A, baculovirus-expressed E1, E2 (UbcH7), and UBE3A were incubated in the presence or absence of ATP and/or E6 at 30°C for 90 min. The reaction mixture was electrophoresed on an SDS-PAGE gel and the proteins visualized by Coomassie

stain. To examine ubiquitination of RING1B by UBE3A, radiolabeled RING1B-I53S was incubated with baculovirus-expressed E1, E2 (UbcH7), and UBE3A at 30°C for 90 min in the presence or absence of ATP and/or E6. The reaction mixture was electrophoresed on an SDS-PAGE gel and visualized by fluorography.

Lentiviral production and transduction of NPCs

HA-tagged E6 was cloned downstream of the ubiquitin C promoter of the FUGW plasmid[100] (Addgene #14883) connected to GFP by T2a (**Supplemental Figure S8A**). pcDNA3 plasmid containing HA-tagged E6[101] and FUGW were used as templates for molecular cloning, which was performed by the iPSC/hESC Genome Targeting division of the UConn-Wesleyan Stem Cell Core. The plasmid was verified by transfecting 293FT cells with E6-containing or normal FUGW using Lipofectamine 3000 (Life Technologies). 48 hours after transfection, GFP expression was confirmed by live cell imaging, and GFP and HA tag expression confirmed by Western blot (**Supplemental Figure S8B**). Lentiviral particles were made by transfecting 293FT cells with 2nd generation packaging systems using Lipofectamine 3000. Virus was concentrated using the Lenti-X Concentrator Kit (Clontech) and viral titer calculated using the qPCR Lentivirus Titration Kit (abm). 48 hours after thawing, 5.5-week NPCs were transduced with FUGW-E6 or FUGW virus for 24 hours by adding concentrated lentivirus directly to NDM. NPCs were plated for terminal differentiation at approximately 6.5 weeks onto glass coverslips. Coverslips were fixed for ICC at 11-11.5 weeks or subject to electrophysiological recordings at 12 weeks.

Quantitative RT-PCR

cDNA was synthesized using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) according to manufacturer's instructions. Quantitative RT-PCR was performed using Taqman Gene Expression Assays and Mastermix (Thermo Fisher Scientific) on the Step One Plus (Thermo Fisher Scientific). Reactions were performed in technical duplicates, with

GAPDH Endogenous Control Taqman Assay used as the housekeeping gene for normalization. Gene expression was quantified using the $\Delta\Delta C_t$ method.

Immunocytochemistry

Immunocytochemistry was performed on stem cells or neurons as previously described.[79] Briefly, cells were fixed using room temperature 4% paraformaldehyde then permeabilized using 0.5% PBS-Triton X 100 (PBS-T) for 10 min at room temperature. Cells were blocked in 0.1% PBS-T containing 2% bovine serum albumin and 5% normal goat serum. Cells were incubated in primary antibody in blocking buffer overnight at 4°C then washed with 0.1% PBS-T. Cells were then incubated in secondary antibody in blocking buffer for 1 hour at room temperature then washed with 0.1% PBS-T. Cells were mounted with either Vectashield Hard Set with DAPI (Vector Laboratories) or ProLong Gold Anti-Fade Hard Set with DAPI (ThermoFisher Scientific) and allowed to set overnight at room temperature before imaging. Slides were imaged on a Zeiss Axiovision microscope at 20x, 40x, or 63x using Axiovision Software (Zeiss). For pluripotency staining of iPSCs/hESCs, the following antibodies were used: rabbit anti-OCT4 (1:200; Stemgent 09-0023), mouse anti-SSEA4 (1:20; ThermoFisher MA1-021), mouse anti-TRA-1-60 (1:200 Santa Cruz sc-21705), anti-mouse Alexa Fluor 488 (1:500). For the E6 experiments, the following antibodies were used: chicken anti-MAP2 (1:10,000, abcam), rabbit anti-HA (1:1200, Cell Signaling), mouse anti-GFP (1:750, Invitrogen), anti-chicken Alexa Fluor 647 (1:250, abcam), anti-rabbit Alexa Fluor 594 (1:400, Invitrogen), anti-mouse Alexa Fluor 488 (1:400, Invitrogen).

Western Blots

Cells were lysed using PathScan Lysis Buffer (Cell Signaling Technology) supplemented with 1 mM phenylmethane sulfonyl fluoride (PMSF; Gibco) and Protease Inhibitor Cocktail III (1:200; Calbiochem). Total protein concentration was quantified using the BCA Protein Assay Kit

(ThermoFisher). Total cell lysate was separated by SDS-PAGE using 4-20% TGX Stain-Free mini gels (BioRad). Protein was transferred to PVDF membrane using the TransBlot Turbo system (BioRad). Membranes were blocked in 5% BSA in TBS-T (Tris-buffered saline plus 0.1% Tween-20) for 1 hour at room temperature then incubated in blocking buffer containing primary antibodies overnight at 4°C. Membranes were washed with TBS-T at room temperature, incubated in blocking buffer containing HRP-conjugated secondary antibodies for 1 hour at room temperature, then washed again in TBS-T. All washes were done three times for 10-12 minutes each at room temperature. Membranes were imaged using the Clarity Western ECL substrate (BioRad) on the ChemiDoc Touch Imaging System (BioRad). The following primary antibodies were used: rabbit anti-UBE3A (Bethyl Laboratories Inc., 1:3000), mouse anti-GAPDH (Millipore, 1:10,000), mouse anti-GFP (Invitrogen, 1:3000), rabbit anti-HA (Cell Signaling Technology, 1:1000). HRP-conjugated secondary antibodies (Cell Signaling Technology) were used at 1:4000 (anti-rabbit).

Statistical Analysis & Graphing

Statistical analysis of all qRT-PCR and electrophysiology data was performed on SPSS Statistics 25 software (IBM). Graphs were generated using Graphpad Prism or R studio.

Live cell imaging

To confirm expression of GFP in live neurons transfected with either E6-FUGW or FUGW virus, coverslips were inverted and placed in a RX-26G imaging chamber (Harvard Apparatus) containing Live Cell Imaging Solution (ThermoFisher). Cells were imaged on the Zeiss 780 confocal microscope with 63x oil immersion objective (numerical aperture of 1.4, 3x zoom). Images were acquired using ZEN microscope software (Zeiss).

Results

Generation of isogenic AS and control stem cell lines

We first sought to generate two isogenic AS and control stem cell pairs. Using isogenic cell pairs would establish an important resource for AS research in which molecular and phenotypic differences caused by normal human variation could be minimized. The first isogenic pair was established by correcting AS Pt Mut iPSCs to a normal genotype using CRISPR/Cas9. These iPSCs were derived from an AS patient with an inherited *UBE3A* missense mutation (**Figure 1A-B**). The missense mutation caused no reduction in the protein levels of UBE3A (**Figure 1C**). This AS-causing mutation was located upstream of a PAM site for *S. pyogenes* Cas9, enabling us to design a CRISPR that preferentially targeted the maternal allele (**Figure 1D**). To correct this mutation, a single-stranded oligonucleotide (ssODN) containing the wildtype UBE3A sequence, and a silent mutation to limit CRISPR cutting of the corrected allele and to facilitate clone screening, was used as the template for homology directed repair (**Figure 1D**). Using this method, we were able to generate a control iPSC line isogenic to the AS iPSCs (Corrected; **Figure 1E-G**).

The second isogenic pair was established by using CRISPR/Cas9 to delete the maternal copy of *UBE3A* in H9 hESCs. For this approach a pair of CRISPRs flanking the UBE3A gene was used to induce double-stranded breaks at the 5' and 3' ends of the gene (**Figure 2A**). Due to error-prone non-homologous end-joining (NHEJ), deletions, duplications, or inversions can occur, which can then be screened for by conventional PCR (**Figure 2A,B**), as previously described.[88] We identified a single clone in which only one copy of *UBE3A* had been deleted and the other allele was devoid of any insertions, deletions, or inversions (**Figure 2B, Supplemental Figure S2**). To determine which copy of *UBE3A* had been deleted, we identified a heterozygous SNP (rs1041933[47]) in an intron of *UBE3A* in H9 hESCs and determined its allelic identity using strand-specific reverse transcription polymerase chain reaction (RT-PCR)

followed by sequencing in neurons (**Figure 2C**). *UBE3A-ATS* is produced exclusively by the paternal allele in neurons and spans the introns of *UBE3A*, which enabled us to infer the parental origin of each SNP allele. We then used Sanger sequencing to determine the identity of the SNP in the genomic DNA of the remaining intact allele. We were able to determine that the maternal allele of *UBE3A* had been deleted (**Figure 2C**). We further confirmed that the other allele had not been disrupted in this clone by assaying *UBE3A* expression by qRT-PCR and Western blot (**Figure 2D-E**), which showed a reduction, but not loss, of UBE3A levels. For both isogenic pairs, it was confirmed that genome editing did not affect expression of pluripotency markers or cause any large-scale genomic rearrangements (**Supplemental Figures S1&S2**).

Isogenic AS stem cell-derived neurons display AS electrical phenotypes

Previously we showed that AS iPSC-derived neurons display impaired neuronal maturation *in vitro*, including a more depolarized resting membrane potential (RMP), less mature action potential (AP) firing patterns, and less frequent spontaneous excitatory synaptic activity.[76] To confirm that our AS neurons also exhibited AS phenotypes when compared to their isogenic controls, we performed whole cell patch clamp recordings. Both AS Pt Mut iPSC-derived and ΔUBE3A$^{m-/p+}$ hESC-derived neurons displayed impaired electrical maturity when compared to their respective isogenic controls (**Figure 3**). This also confirmed that the genome editing of the stem cells achieved its desired effects: the correction of the AS mutation in the iPSCs produced electrically normal neurons (**Figure 3A-C)**, while deletion of *UBE3A* in normal hESCs produced neurons that recapitulated most AS phenotypes (**Figure 3D-F**).

Establishing a molecular AS phenotype using the transcriptome

Currently there are multiple approaches being explored as potential AS therapies, many of which are based on DNA sequences unique to humans. Testing these therapies in mouse requires either use of a different sequence, which does not actually test the putative human

therapeutic, or the costly and at times challenging development of humanized murine models. Stem cell-derived neurons allow us to test these potential therapies on human neurons in dish, a step that is especially necessary for nucleic acid-based therapeutic approaches. Just as important as being able to test these drugs, however, is the ability to begin to assess their efficacy. We wanted to see if we could use the transcriptome of human neurons to establish a quantitative molecular phenotype for AS. Specifically, we sought to identify a list of genes that are consistently differentially expressed between AS and neurotypical stem cell-derived neurons. Our goal in this approach was not identify putative UBE3A targets that might be affected in AS but to get a definitive set of genes that reflect our ability to restore function to AS neurons following drug treatment.

mRNAseq was performed on 12-week neurons differentiated from the AS Pt Mut, Corrected, H9, and ΔUBE3A$^{m-/p+}$ stem cell lines. Three separate analyses were conducted to produce a conservative list of genes consistently differentially expressed between AS and neurotypical neurons: AS Pt Mut vs Corrected, H9 vs ΔUBE3A$^{m-/p+}$, and AS+ΔUBE3A$^{m-/p+}$ vs Corrected+H9. For each of the three comparisons, surrogate variable analysis was performed prior to intersection with the other pairs to remove the variation not conferred by genotype (see **Supplemental Figure S4**). The final list of differentially expressed (DE) genes consisted of genes that were significantly differentially expressed in all three analyses (n = 855 genes, padj < 0.01; **Figure 4A-B**). As we suspected, this gene list did not contain any known ubiquitylation substrates for the *UBE3A* protein and most likely reflects genes whose change in expression were secondary or tertiary to loss of UBE3A function. Indeed, when we examined gene expression changes in normal hESC-derived neurons which had undergone an acute antisense oligonucleotide-mediated *UBE3A* knockdown, we saw significant changes in 15 genes, none of which, aside from UBE3A itself, had dramatic changes in gene expression (**Supplemental Figure S6**). This suggests that there are few, if any, direct transcriptional targets of UBE3A in our human neuron cultures. Furthermore, of the small handful of significant genes, none

appeared on our list of AS DE genes. Finally, gene ontology (GO) analysis of our set of AS DE genes revealed differences in processes such as extracellular matrix, cell adhesion, and synaptic signaling (**Figure 4C**).

Rescue of AS differentially expressed genes by restoring UBE3A function

To determine whether genes identified by mRNAseq could be used as a readout for treated AS neurons, we examined the expression of 11 genes from our list by qRT-PCR. We wanted to assess whether we could see restoration of the transcriptome phenotype after restoring function to the mutant *UBE3A* protein in AS point mutation neural progenitor cells (NPCs). This missense mutation in the AS Pt Mut iPSCs changes an amino acid in the HECT domain of the UBE3A protein, which confers ubiquitin ligase activity. Because this mutation did not cause a loss of UBE3A protein (**Figure1C, G**), we hypothesized that it might instead cause a loss of protein function. As seen in **Figure 5A**, the ability of the mutant version of UBE3A to ubiquitinate itself or RING1B was severely diminished in an *in vitro* ubiquitination assay. Interestingly, in the presence of the E6 oncoprotein, ubiquitin ligase activity of the mutant version of UBE3A could be rescued. In the context of cervical cancer, the E6 oncoprotein, produced by the human papilloma virus (HPV), acts as an allosteric activator of UBE3A, allowing it to ubiquitinate substrates that it normally would not target, such as p53.[41]

We reasoned that E6 might also be able to activate this mutant version of UBE3A in our cultured AS patient-derived cells. Five-week AS Pt Mut neural progenitor cells were treated with virus expressing HA-tagged E6 and GFP or virus expressing GFP alone (**Figure 5B; Supplemental Figure S7C**). Whole-cell patch clamp recordings of GFP+ cells revealed that treatment with E6 rescued resting membrane potential in AS Pt Mut neurons while having no effect on RMP in the isogenic control E6-treated neurons (**Figure 5C**). Conversely, E6 treatment did not significantly affect AP firing in AS Pt Mut neurons, but appeared to negatively affect AP firing in Corrected neurons (**Figure 5C**), indicating that E6 expression may be detrimental to

some aspects of normal neuronal function. We next assayed the expression of 11 genes from the DE gene list that had an absolute log2 fold change (log2fc) of greater than one. All 11 genes assayed recapitulated the gene expression changes seen in the mRNAseq data (**Figure 5D (black)**) when comparing differences in expression between the AS Pt Mut/GFP to the Corrected/GFP neurons. Importantly, treatment of AS neurons with E6 caused significant changes in gene expression in seven of the genes assayed (**Figure 5D (red)**). Together, these results validate our mRNAseq data and demonstrate the utility of using the transcriptome as a phenotype for AS neurons following treatments.

Discussion

AS is a rare neurodevelopmental disorder for which there is currently no cure, that is caused by loss of the maternal copy of the *UBE3A* gene. Previously, we had established AS iPSC lines, which were shown to undergo normal imprinting of *UBE3A* upon differentiation into neurons.[78] We also showed that AS iPSC-derived neurons display impaired neuronal maturation compared to control iPSC-derived neurons, and that these electrophysiological phenotypes were a result of the lack of *UBE3A* in these cells.[76] Here, we have used isogenic AS and neurotypical neurons to establish a robust quantitative molecular phenotype for AS neurons by comparing the transcriptomes of isogenic AS and neurotypical neurons We have identified a list of genes that can be assayed in human neuronal cultures following drug treatments to determine the ability of the drug to rescue human AS stem cell derived neurons.

We chose to employ a more conservative approach to our mRNAseq analyses in order to establish a robust and replicable phenotype for AS neurons. First, we established two isogenic pairs of AS and control stem cells, one by correcting AS in iPSCs and one by causing AS in hESCs. By examining more than one pair of isogenic cells, which were derived from individuals with two distinct genetic backgrounds, we can more confidently say that the gene list identified reflects gene expression changes seen in AS neurons. Second, we performed a third analysis in which all 12 AS samples were compared to the 12 control samples, and looked for genes in common between this analysis and the other two analyses. Finally, we employed surrogate variable analysis[98] in an attempt to account for other extraneous factors[102] that could have caused differences in gene expression, which has been previously shown to improve the accuracy of fold-change estimates[103]. The result of this more cautious approach was still a relatively large list of 855 genes that were differentially expressed in AS versus neurotypical neurons. We propose that this list of genes, as a whole or in part, can be assayed following drug

treatments of AS neurons by any desired means, such as qRT-PCR, custom qRT-PCR array, targeted sequencing, or whole transcriptome profiling.

To validate the use of the transcriptome as a phenotype, we assayed a small portion of this gene list by qRT-PCR after restoring ubiquitin ligase function to UBE3A in the AS Pt Mut NPCs. AS Pt Mut and Corrected NPCs were treated with virus expressing E6 oncoprotein, which we showed can rescue the activity of the mutant version of UBE3A seen in this AS patient in an *in vitro* ubiquitination assay. We showed that treatment of AS NPCs with E6, in addition to rescuing resting membrane potential deficits in AS neurons, was able to partially rescue the expression of the DE genes assayed. While treatment with E6 itself is not an ideal potential AS therapy due to potential off-target gene expression changes independent of UBE3A function, these findings confirm the robustness and utility of using differentially expressed AS genes as a readout for *in vitro* drug treatments. This phenotype will be especially useful for therapies that specifically modulate UBE3A's expression or function, such as through unsilencing of paternal *UBE3A* or through introduction of exogenous *UBE3A* expression.

The results of these experiments were not intended to identify putative pathways that are misregulated in AS, nor were they intended to identify putative transcriptional or ubiquitylation targets of UBE3A in neurons. The gene expression changes seen in the AS neurons instead reflect secondary or tertiary changes following loss of UBE3A and, presumably, overabundance of UBE3A's as-yet-unknown neuronal targets. Consistent with this notion, gene expression changes seen in neurons with acute UBE3A knockdown by ASO, which would more accurately reflect primary gene expression changes immediately following UBE3A loss, do not reflect the changes seen in the AS neuron transcriptome, and overall very few genes are changed. That being said, some of the differentially expressed genes do recapitulate physiological AS *in vitro* phenotypes and may suggest other important AS pathologies. For example, we see decreased expression of mature neuron markers such as *TBR1*, *NEUROD2*, and *BCL11B* (CTIP2) in AS neurons compared to neurotypical controls, and GO analysis revealed decreases in genes

implicated in synaptic transmission. These results are reminiscent of the decreased electrical maturity in AS neurons that we have shown previously[76] and in this study. Additionally, we see an upregulation of genes involved in the extracellular matrix and in cell adhesion in AS neurons compared to controls, which may indicate a novel pathology for cultured AS neurons. Interestingly, we showed that we could reactivate the mutant form of UBE3A by introduction of the E6 oncoprotein, a known allosteric activator of UBE3A,[41] and rescue both electrical and transcriptome phenotypes. While E6 itself is not a viable therapeutic target for AS, these data serve as an important proof of principle that for some AS-causing mutations, we can rescue or increase UBE3A's ubiquitin ligase activity through its activation. It is possible that the effects of E6 on UBE3A could be mimicked by small molecules, which could be a therapeutic approach for a subset of AS patients.

In addition to our transcriptome phenotype, we have also demonstrated that the AS Pt Mut and ΔUBE3A$^{m-/p+}$ stem cell-derived neurons display the same electrical phenotypes demonstrated previously. This is important because it demonstrates that AS neurons display these electrical phenotypes regardless of which protocol is used to differentiate them (monolayer, shown here, versus embryoid body, used previously). Additionally, both the electrophysiological and transcriptome phenotypes are seen in neurons derived from the AS Pt Mut iPSCs, which actually have normal levels of UBE3A protein, but produce a version of the protein that is virtually ligase-dead. This implies that loss of UBE3A's ubiquitin ligase activity alone is sufficient to produce AS, both *in vitro* and *in vivo* (as demonstrated by the two AS patients with this mutation). Although we did not see significant changes in the frequency of spontaneous excitatory synaptic activity in the ΔUBE3A$^{m-/p+}$ neurons when compared to their isogenic control, we believe that this may be due to the decreased number of cells used for these recordings as compared to the AS Pt Mut and Corrected neurons, and also compared to our previous work. Supportive of this notion is the fact that the difference between the AS Pt Mut and Corrected neurons (.27 Hz) is of similar magnitude to the difference between the H9 and

ΔUBE3A$^{m-/p+}$ neurons (.45 Hz). It should also be noted, however, that these neurons in general had dramatically more frequent spontaneous excitatory activity than the AS and corrected neurons and also in our previous work, so it is possible that AS neurons with a higher frequency of spontaneous synaptic events do not display these deficits.

This study is not without its limitations. AS is caused by large deletions of the 15q11-q13 region in the majority of patients.[22] These deletions encompass not only *UBE3A*, but multiple other non-imprinted genes, such as *GABRB3, GABRA5,* and *HERC2*. It was our goal to identify gene expression changes in AS neurons that arise specifically due to loss of *UBE3A*, which we can ideally apply to neurons from AS patients with any type of genetic etiology. It is possible, however, that reduced expression of the non-imprinted genes in this chromosomal region could affect expression of some of the genes from our list in cells derived from these patients. Another caveat is that our conservative criteria for establishing our list of AS DE genes may have excluded other AS DE genes, which could have impacted the results of our GO analysis. Since our goal here was not to identify or study pathways altered in AS, we were not concerned by this. However, it is something that should be kept in mind when considering the GO results presented here.

In summary, we have demonstrated that AS neurons, when compared to isogenic controls, have changes in the transcriptome that are consistent across isogenic pairs. We have identified a list of genes that can serve as a robust molecular phenotype for cultured human stem cell-derived AS neurons, which we can now assay in rescue experiments. We propose to use this phenotype as high-throughput method for assessing the efficacy of potential AS therapeutics in cultured human neurons. This is a necessary step, especially for gene therapy-based approaches, before these therapies can proceed to clinical trials. By more thoroughly vetting potential therapies before trials, we can improve the likelihood of identifying a therapy for AS that is both safe and effective.

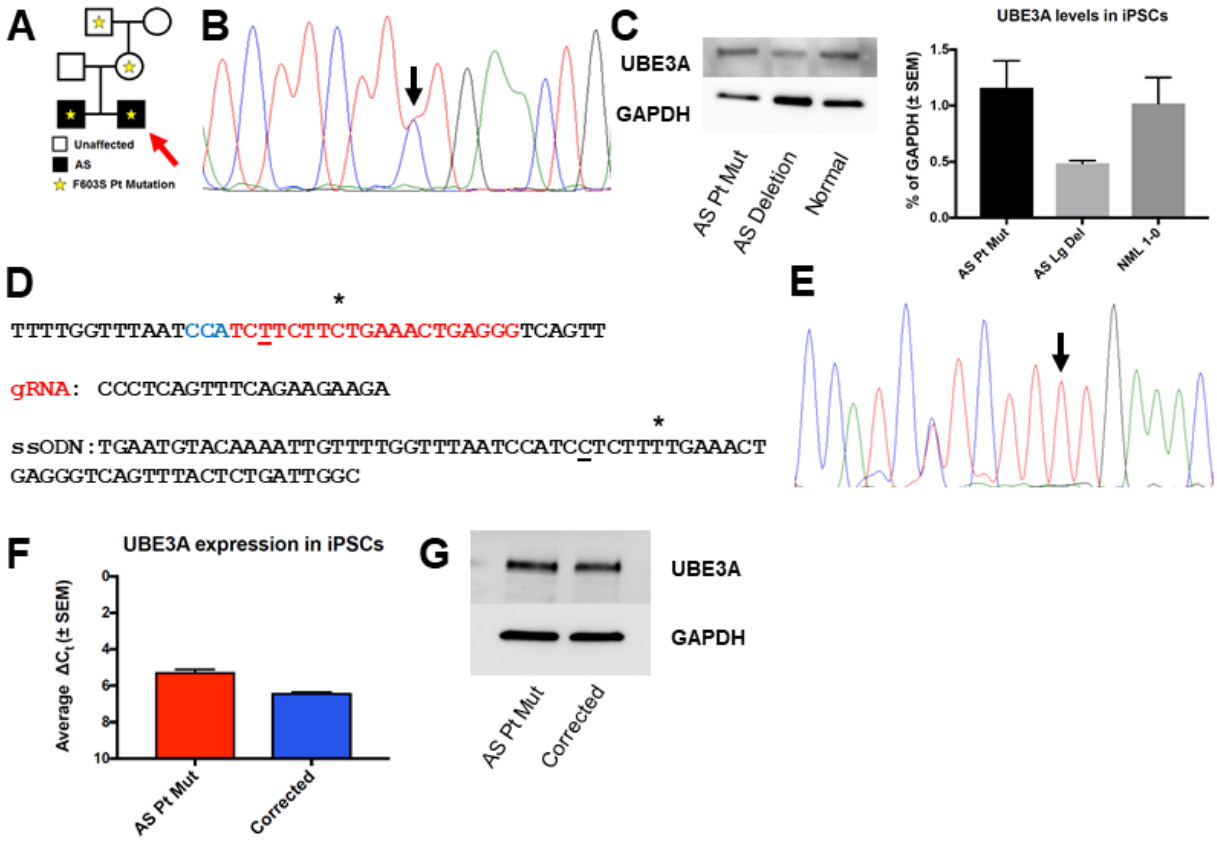


Figure 1. Generation of isogenic control iPSC line by correcting AS point mutation
A Pedigree showing family with AS point mutation. Arrow indicates proband. Star indicates AS point mutation. **B** Sanger sequencing of iPSC gDNA showing point mutation (arrow) **C** Western blot showing normal levels of UBE3A protein in AS Pt Mut iPSCs compared to AS large deletion (AS Del 1-0) and normal (NML1-0) iPSCs **D** Top: Location of gRNA binding site (red), PAM (blue), and CRISPR cut site (T) Middle: gRNA sequence; Bottom: ssODN sequence * = location of mutation/correction; C = silent mutation to create FokI site **E** Sanger sequencing of iPSC gDNA showing correction of point mutation following genome editing **F** qRT-PCR showing *UBE3A* levels in AS and corrected iPSCs **G** Western blot showing UBE3A levels in AS Pt Mut and Corrected iPSCs

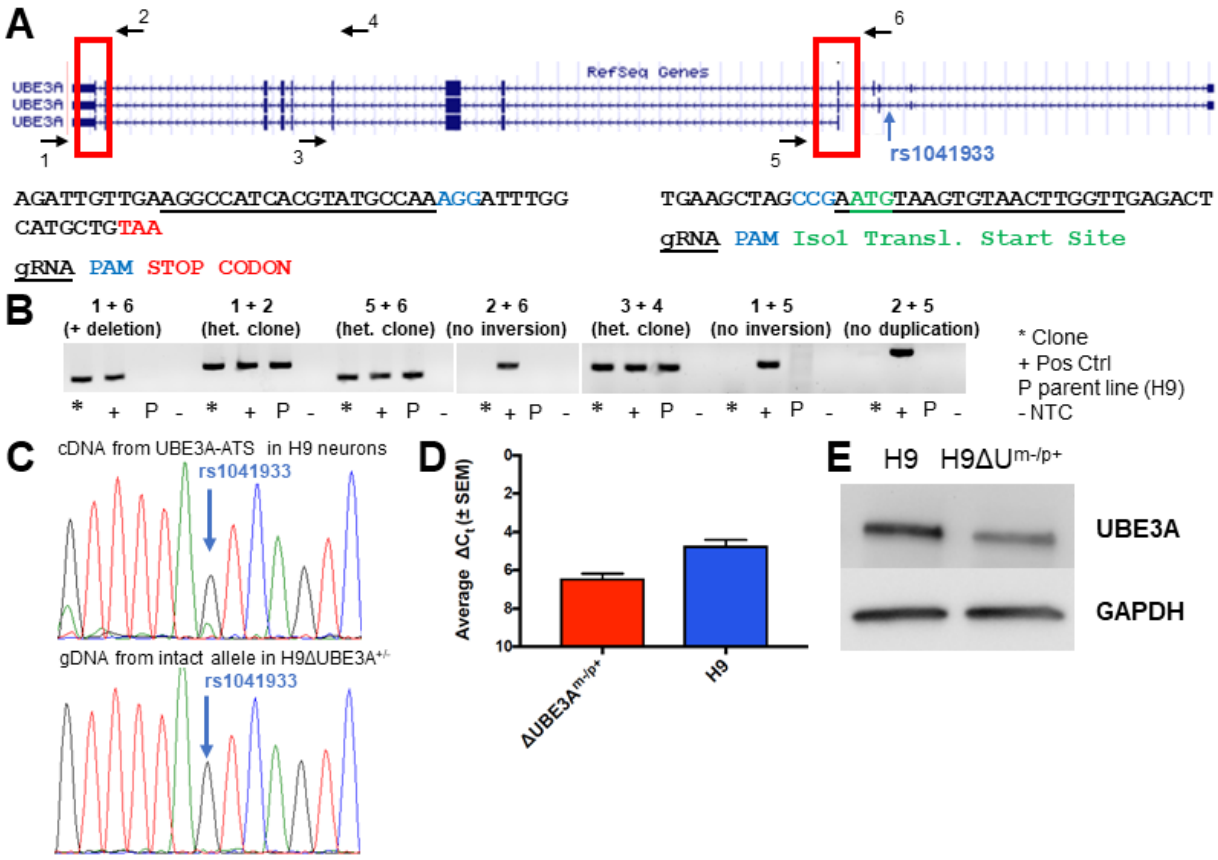


Figure 2. Generation of isogenic AS hESCs by deleting maternal UBE3A
A Top: Schematic illustrating CRISPR gRNA and primer locations, as well as the location of rs1041933 (light blue arrow). Red box indicates approximate location of gRNA binding/CRISPR cutting. Bottom: Location of gRNA (underlined) and PAM site (blue), as well as the STOP codon (red) and translational start site (green). **B** PCR of gDNA indicating heterozygous deletion of one copy of UBE3A. Numbers correspond to primers in part A above. **C** Top: Sanger sequencing of cDNA from the UBE3A-ATS transcript. Bottom: Sanger sequencing of gDNA from the intact allele of the deletion clone showing SNP rs1041933 **D** qRT-PCR showing UBE3A levels in H9 and ΔUBE3A$^{m-/p+}$ hESCs **E** Western blot showing UBE3A levels in H9 and ΔUBE3A$^{m-/p+}$ hESCs

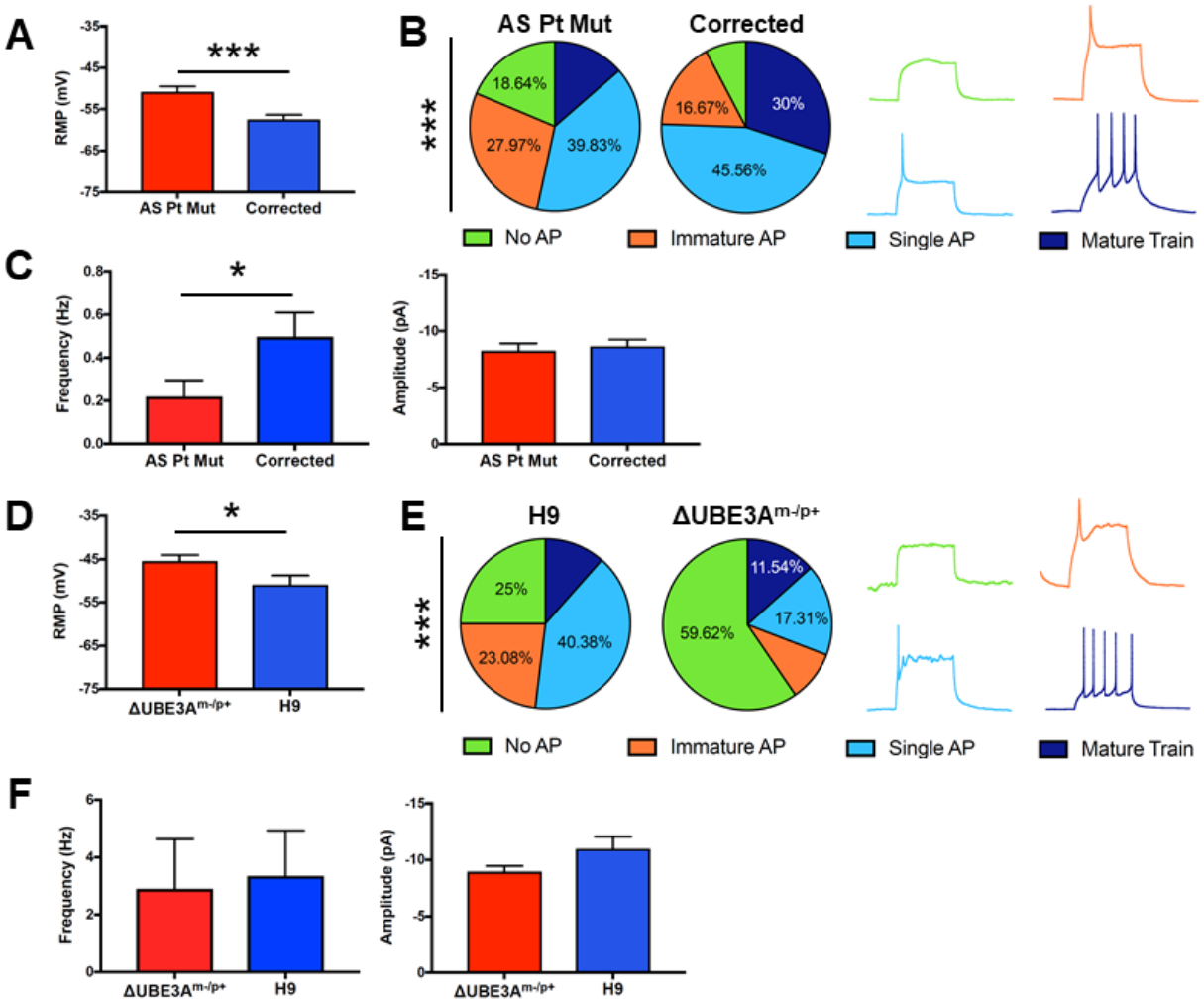


Figure 3. Isogenic AS neurons display AS electrical phenotypes
A Resting membrane potential in AS Pt Mut and Corrected iPSC-derived neurons **B** Left: action potential firing characterization of AS Pt Mut and Corrected neurons; right: sample AP traces **C** Excitatory spontaneous synaptic activity frequency (left) and amplitude (right) in AS Pt Mut and Corrected neurons. n = 90-120 neurons on 6-8 coverslips per cell line **D** Resting membrane potential in ΔUBE3A$^{m-/p+}$ and H9 hESC-derived neurons **E** Left: AP firing characterization of ΔUBE3A$^{m-/p+}$ and H9 neurons; right: sample AP traces **F** Excitatory spontaneous synaptic activity frequency (left) and amplitude (right) in ΔUBE3A$^{m-/p+}$ and H9 neurons. n = 53-54 neurons on 4 coverslips per cell line. * $p < 0.05$; *** $p < 0.005$; RMP, frequency, amplitude: student's t-test; AP firing: X^2 test

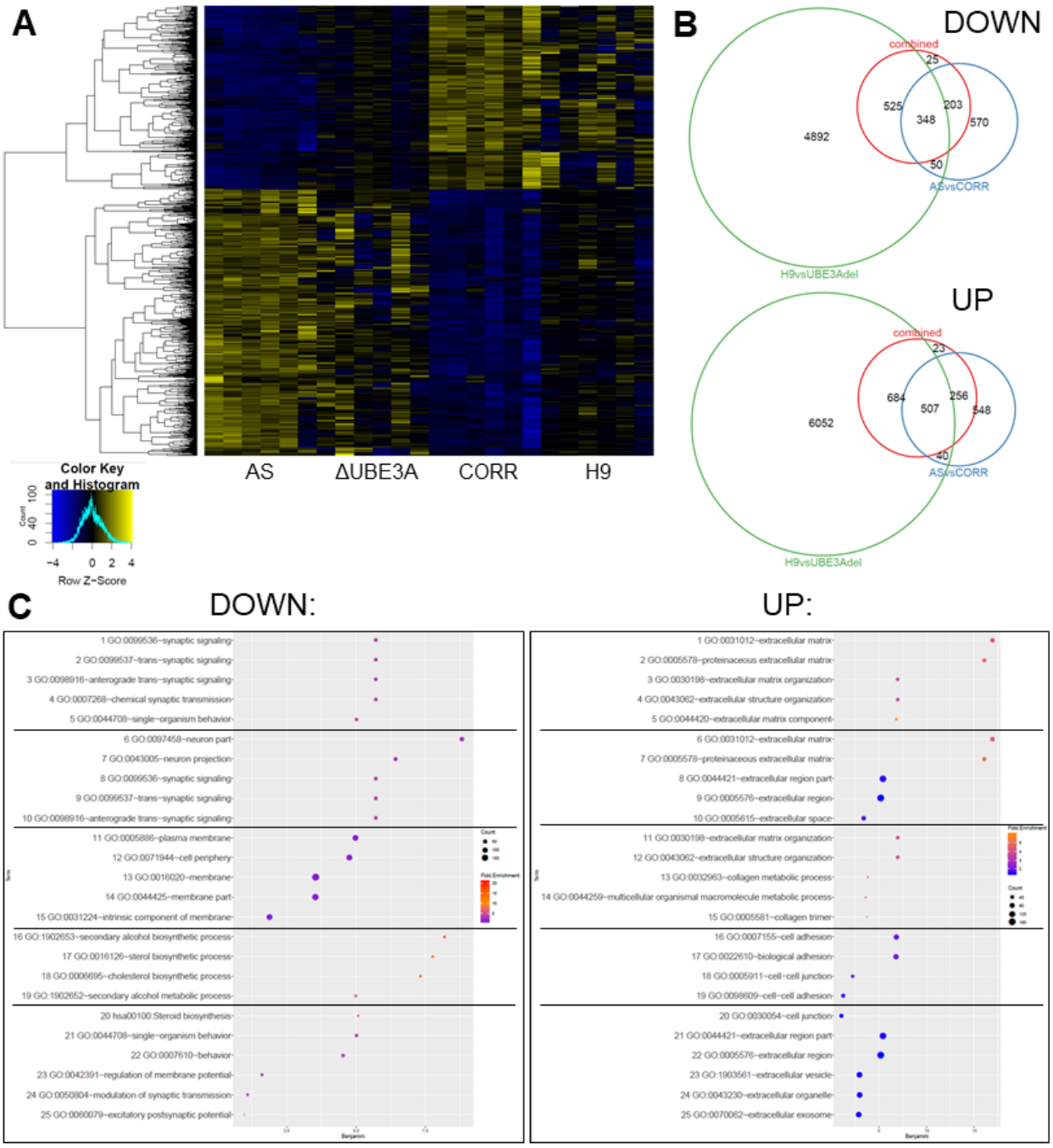


Figure 4. Establishing a transcriptome phenotype for AS neurons
A Heatmap showing expression of 855 genes differentially expressed in AS (padj < 0.01) **B** Weighted Venn diagrams comparing the three mRNAseq analyses showing common genes decreased (top) or increased (bottom) in AS neurons **C** Dot plots showing the top 5 significant functional annotation clusters and the top 5 GO terms in each cluster for DOWN (left) and UP (right) AS DE genes. Dot size indicates number of genes belonging to that term, color indicates enrichment score. Benjamini test p < 0.05

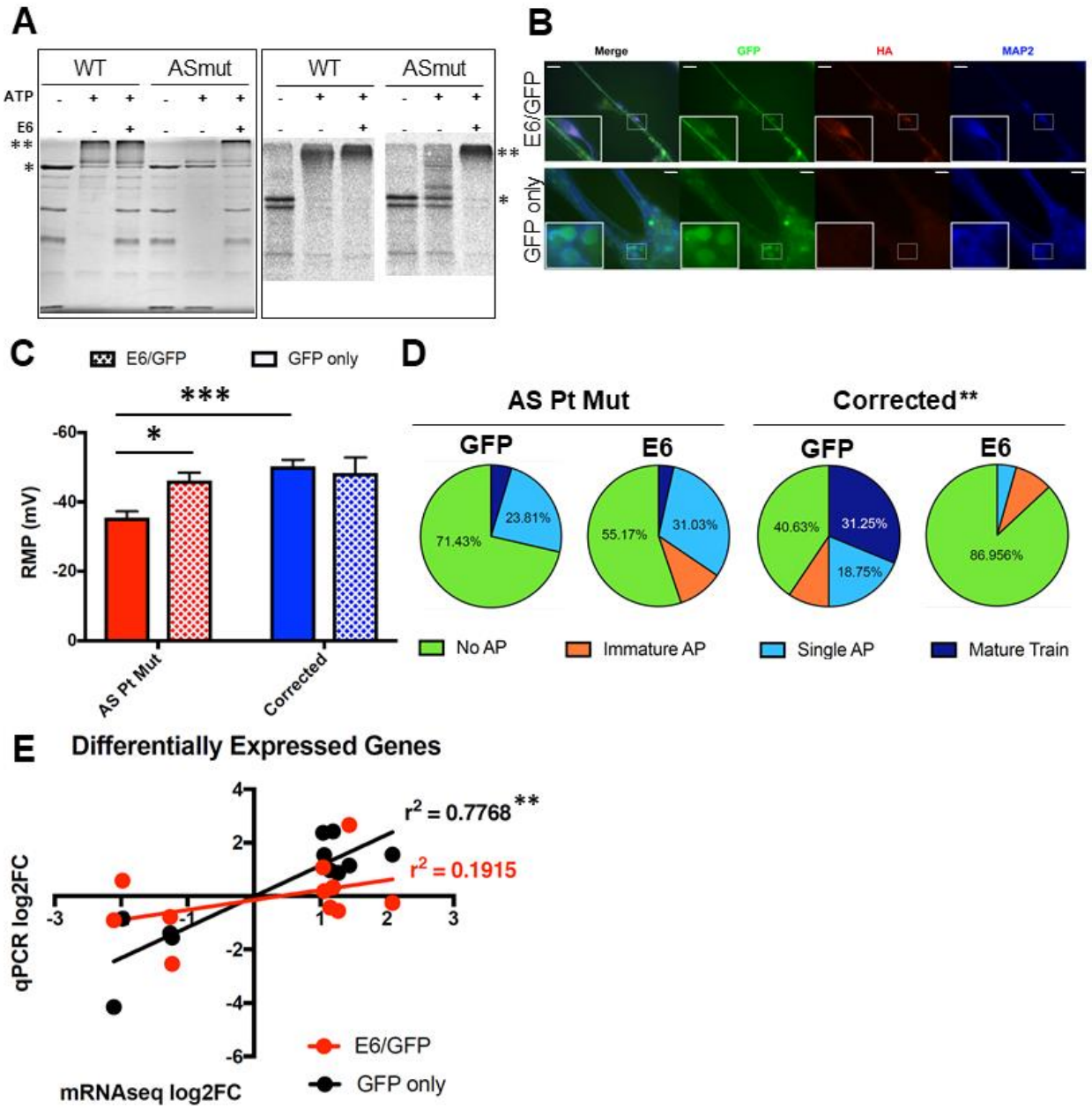


Figure 5. Rescue of ubiquitin ligase activity of mutant UBE3A partially restores AS phenotypes

A *In vitro* ubiquitination assay showing levels of ubiquitinated (**) and non-ubiquitinated (*) UBE3A (left) or RING1B (right) in the presence of wild-type (WT) or mutant (AS mut) UBE3A protein **B** Immunocytochemistry showing expression of GFP (green) in virus-treated MAP2+ (blue) neurons and expression of HA-tagged E6 (red) in E6-treated neurons **C/D** Effect of E6 treatment on resting membrane potential (RMP) and action potential firing (AP firing) AS phenotypes. n = 23-34 neurons on 6-8 coverslips per condition. **E** Scatterplot showing correlation between expression of DE genes from mRNAseq data (x-axis) and qRT-PCR of E6 (red) or GFP (black) treated neurons (y-axis). Scale bar = 20 μm; * $p < 0.05$ ** $p < 0.01$ *** $p < 0.005$ (C, students t-test; D, X^2 test; E, linear regression)

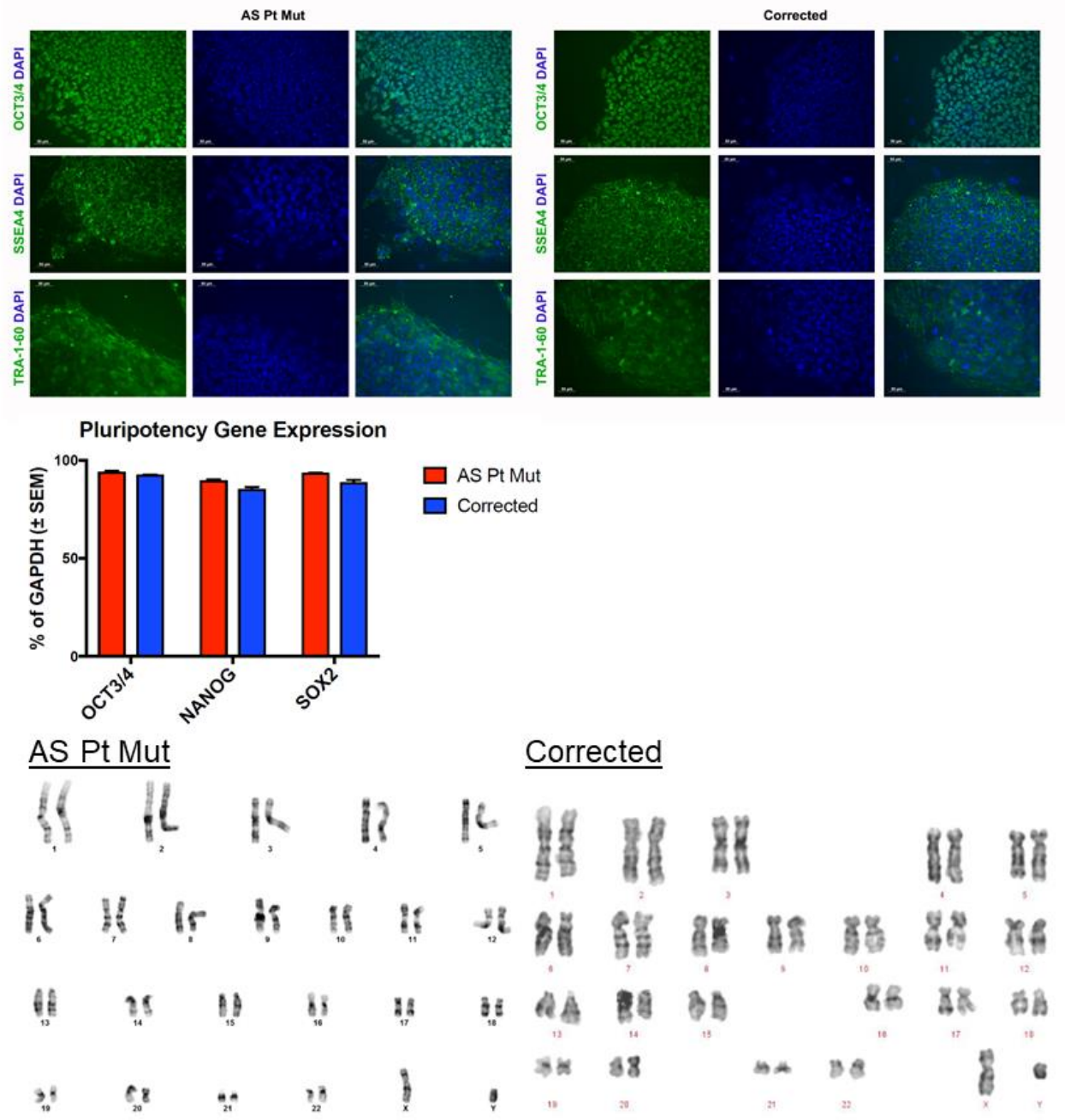


Figure S1. Characterization of AS Pt Mut and Corrected iPSCs

Top: expression of pluripotency markers in AS Pt Mut (left) and Corrected (right) iPSCs
Middle: qRT-PCR showing that AS Pt Mut and Corrected iPSCs have similar levels of expression of pluripotency genes
Bottom: AS Pt Mut (left) and Corrected (right) iPSCs have a normal karyotype

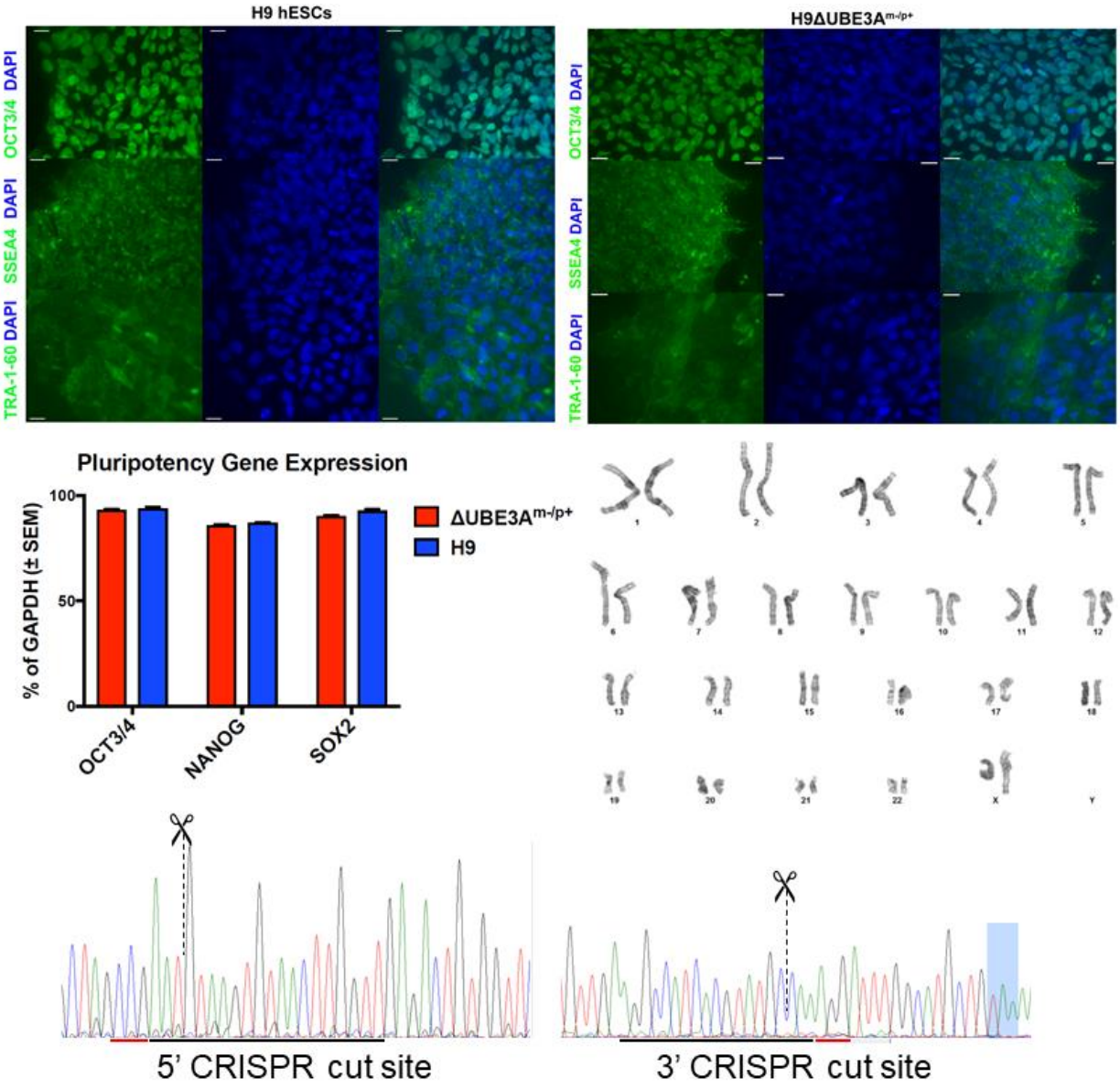


Figure S2. Characterization of ΔUBE3A$^{m-/p+}$ hESCs

Top: CRISPR editing of H9 line did not affect expression of pluripotency markers
Middle left: qRT-PCR showing that ΔUBE3A$^{m-/p+}$ hESCs express similar levels of pluripotency genes compared to its parent line, H9
Middle right: ΔUBE3A$^{m-/p+}$ hESCS have a normal karyotype after genome editing
Bottom: Sanger sequencing shows no indels present at the CRISPR cut sites at either the 5' (left) or 3' (right) sites. Scissors indicates the specific CRISPR cut site. Black line indicates sgRNA binding site. Red line indicates the PAM site. The UBE3A STOP codon is highlighted in blue (right).

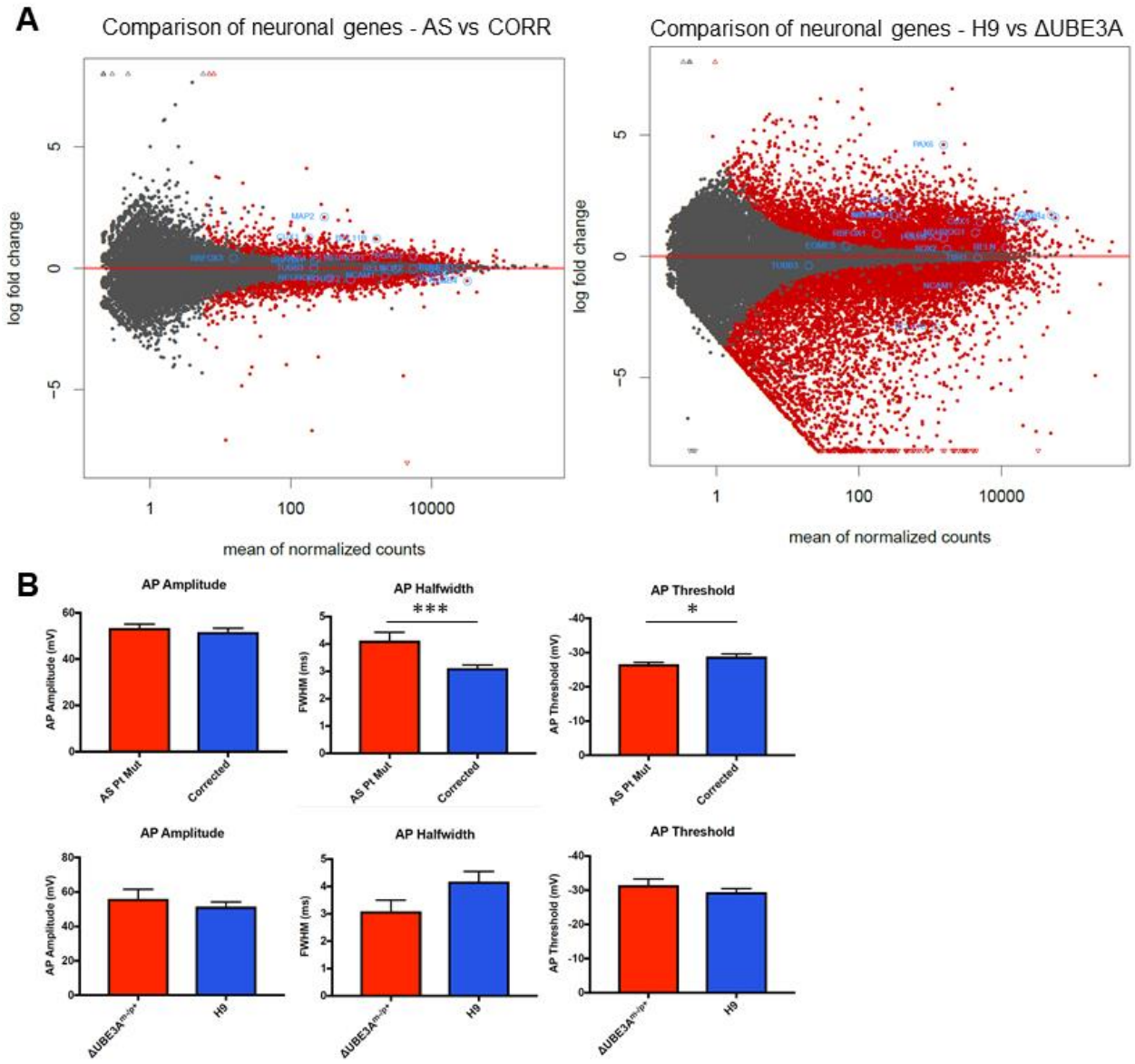


Figure S3. Additional information about the hPSC-derived neurons

A MA plots showing DE genes (red) in AS Pt Mut vs Corrected (left) and H9 vs ΔUBE3A$^{m-/p+}$ (right) neurons. Neuronal and neural development genes are highlighted in blue. **B** Action potential properties in AS Pt Mut vs Corrected (top) and H9 vs ΔUBE3A$^{m-/p+}$ (bottom) neurons. * $p < 0.05$, *** $p < 0.005$ (student's t-test)

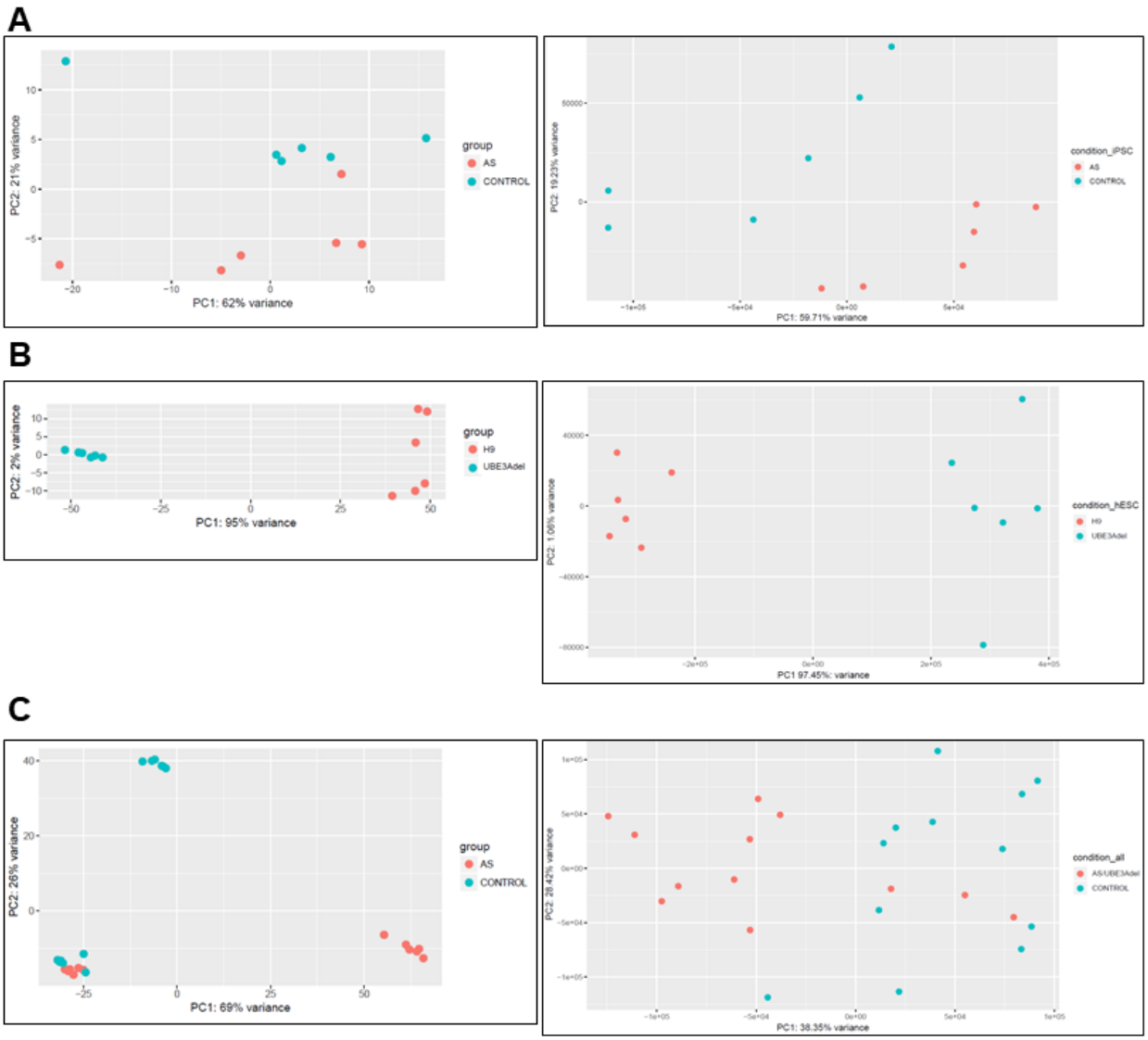


Figure S4. Principal component analysis plots before and after surrogate variable analysis

Left: PCA before SVA, right: PCA after SVA. **A** AS Pt Mut vs Corrected **B** H9 vs ΔUBE3A$^{m-/p+}$ **C** Combined analysis

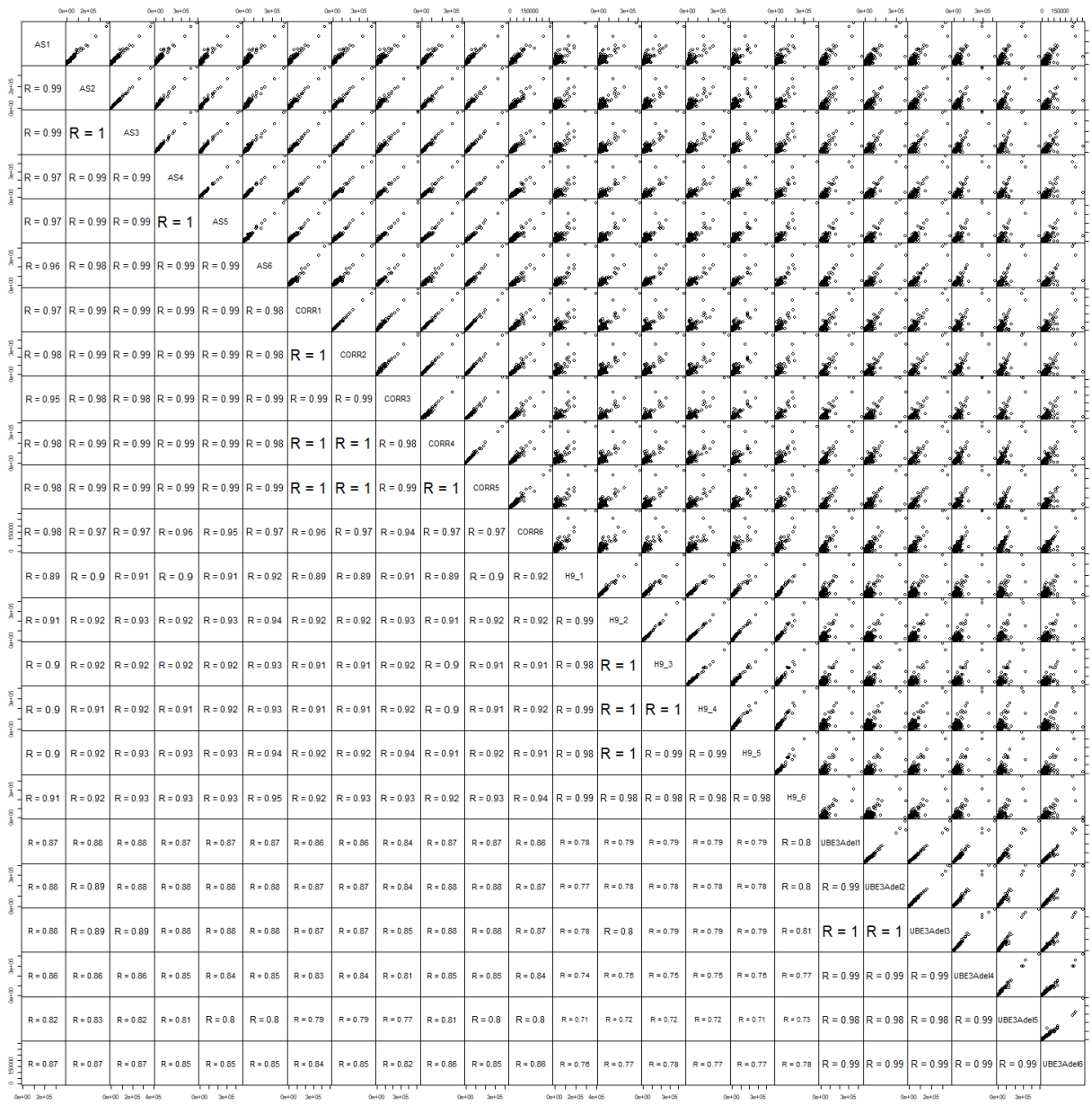


Figure S5. Scatterplot matrix

Scatterplot matrix showing degree of correlation between all 24 samples used in mRNAseq analysis to establish transcriptome phenotype. R indicates Pearson correlation.

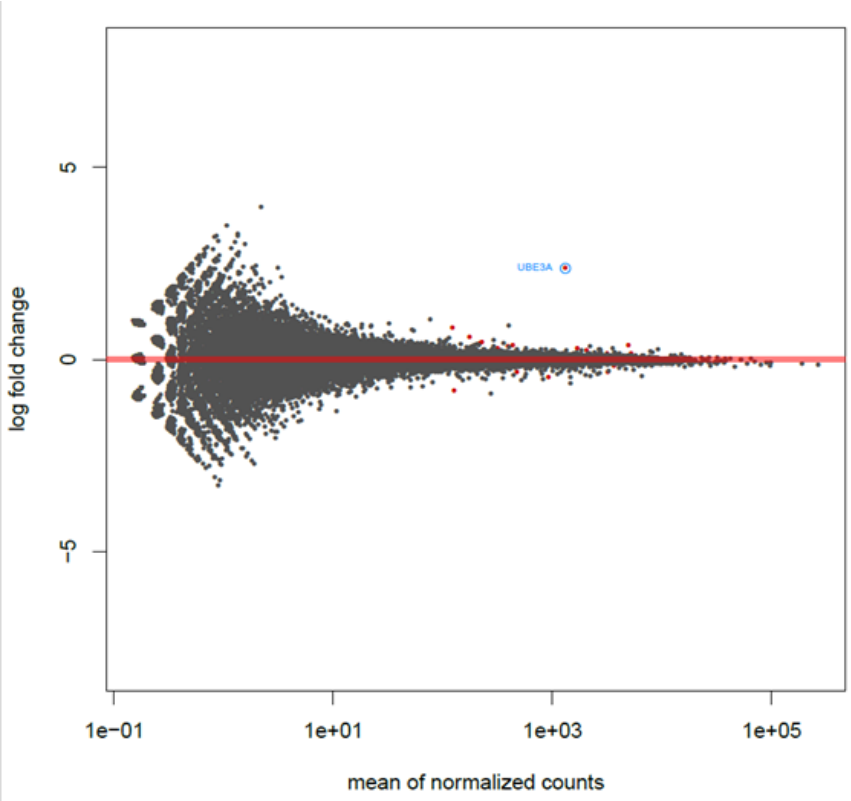


Figure S6. DE genes after acute UBE3A knockdown via ASO treatment

MA plot showing DE genes (red) in ASO- vs scramble-treated H9 neurons. UBE3A is highlighted in blue.

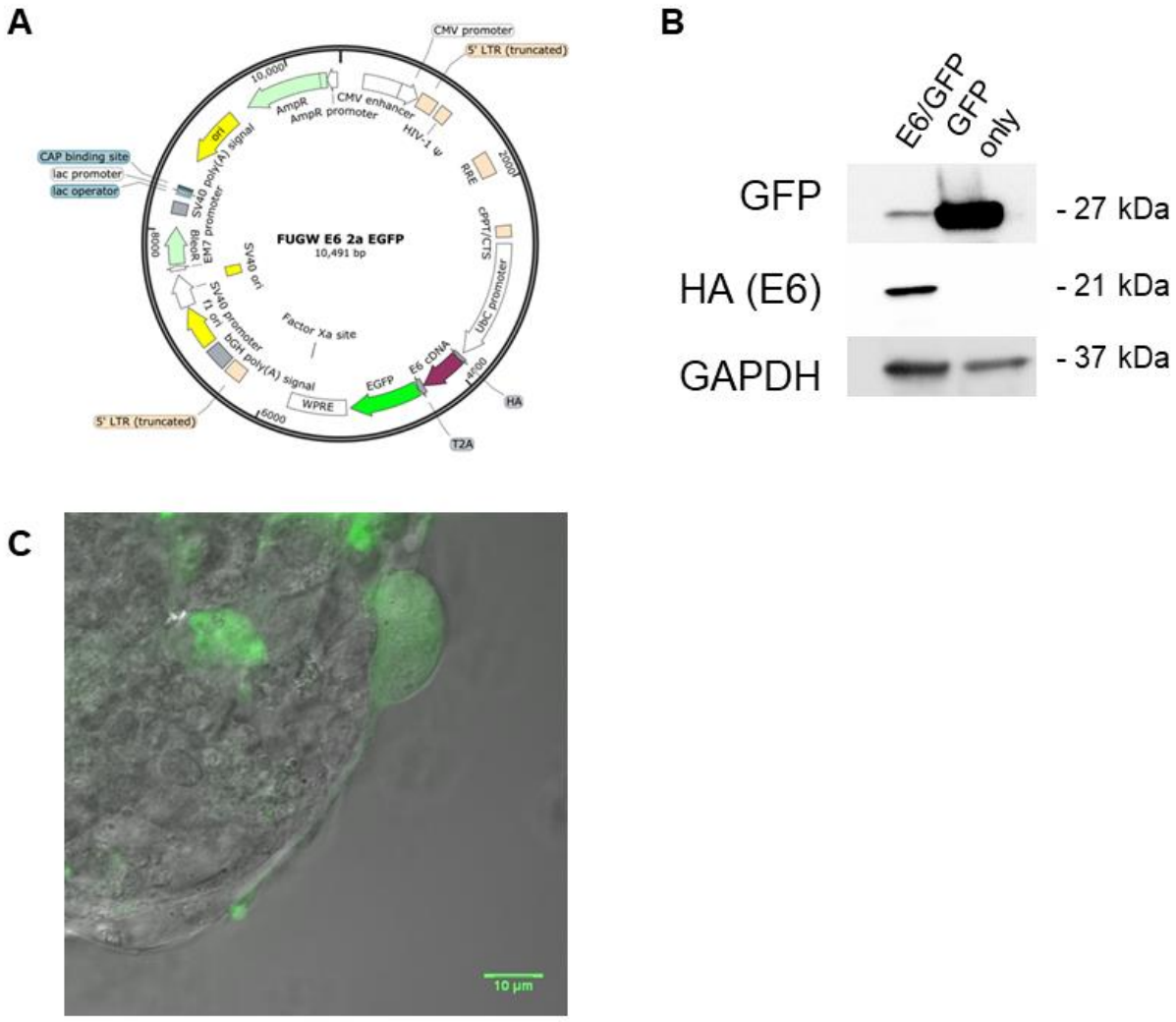


Figure S7. Generation and validation of E6-GFP system

A Map of FUGW vector after insertion of HA/E6-T2A-GFP **B** Western blot showing expression of GFP and HA tag in HEK293T cells following transfection with E6-FUGW or FUGW vector. **C** Live cell image of FUGW transfected neurons showing transfection of only a proportion of neurons. Scale bar = 10 µm

Table S1 PCR Primers

Name	Sequence
Point Mut F	TTCATCAGTCCTTAATAAAATACAAAA
Point Mut R	TTTTCCCCATTAGCTTCCT
Point Mut Seq	AGGTATGTTCACATACGATG
Iso1 F	GCTTATAATGGCTTGTCTGTTGG
Iso1 R	CTGCTACCAGGGAAGCAAAA
Iso1 Seq	TTCTTTCATGTTGACATCTTTAATTTT
UBE3A 3' R	TGGGACACTATCACCACCAA
UBE3A_del_F1	AGTCCAACCCTTAAAATAAATGTG
UBE3A_del_R1	CCCACATGTCCCCAATAAAG
UBE3A_intron12_F1	GGCAACTTGGTAGTTACACAACA
UBE3A_intron12_F2	CCCATGACTTACAGTTTTCCTG
EndogNeo F	TGAAACACTTTGGAAATGTAGCC
EndogNeo R	GCACTTGAGAAAACAATGTCCA
SNP -933 F	TTAAGCAGTTGCCCTCCTTG
SNP -933 R	AAAATTAAGCAGCCTCCGAGT
SNP -933 Seq	TGACTGAATTTGTCCGTATTTGA
SDM 1F	ATGCCATTGTTGCTGCTTC
SDM 1R new	CAAGTGTATGCTACTTAGATGTTTTAACAAAACCAAAT TAGGTAGAAGAAGACTTTGACT
SDM 1R ctrl	CAGAAGAAGATGGATTAAACCAAAA
SDM 2F new	CCATCTTCTTCTGAAACTGAGGGTCAGTTTACTCTGAT TGGCATAGTACTGGGTCTGGCT
SDM 2R	CCTGTAGACAACCATGGGAAA
SDM seq F	ATGCAGACCAGATTCGGAGA
SDM seq R	TCATCTTCCACATTCCCTTCA

Table S2 Taqman assays used for mRNAseq validation

Gene	Assay
NEUROD2	Hs00272055_s1
COL6A6	Hs01029204_m1
GRM3	Hs00932301_m1
TBX18	Hs01385457_m1
GJA5	Hs00979198_m1
WISP1	Hs04234730_m1
BCL11B	Hs01102259_m1
MGP	Hs00969490_m1
TBR1	Hs00232429_m1
EDNRA	Hs03988672_m1
C7	Hs00175109_m1

Table S3 mRNAseq sample information

Sample	Trimmomatic	STAR			
	% passing - paired	Total mapped	Uniquely mapped	Multi-mapped	Unmapped
AS1	95.79%	78.60%	74.82%	3.78%	21.40%
AS2	93.93%	94.87%	90.55%	4.32%	5.13%
AS3	93.44%	88.38%	84.40%	3.99%	11.62%
AS4	99.87%	94.40%	90.37%	4.03%	5.60%
AS5	99.29%	93.56%	89.51%	4.05%	6.44%
AS6	98.58%	88.23%	84.24%	3.98%	11.77%
CORR1	99.89%	93.88%	89.71%	4.17%	6.12%
CORR2	99.41%	89.03%	84.79%	4.25%	10.97%
CORR3	98.51%	90.58%	86.13%	4.44%	9.42%
CORR4	99.85%	96.57%	92.36%	4.21%	3.43%
CORR5	99.92%	92.57%	88.44%	4.13%	7.43%
CORR6	97.98%	88.50%	84.39%	4.11%	11.50%
H9 1	99.82%	95.89%	92.17%	3.72%	4.11%
H9 2	99.90%	89.27%	85.97%	3.30%	10.73%
H9 3	99.88%	94.86%	91.15%	3.71%	5.14%
H9 4	99.82%	93.46%	89.89%	3.57%	6.54%
H9 5	99.86%	92.67%	89.05%	3.62%	7.33%
H9 6	99.83%	94.58%	90.02%	4.55%	5.42%
UBE3Adel 1	99.92%	93.86%	89.89%	3.97%	6.14%
UBE3Adel 2	99.95%	93.66%	89.27%	4.39%	6.34%
UBE3Adel 3	99.94%	89.08%	85.16%	3.93%	10.92%
UBE3Adel 4	99.96%	94.23%	86.09%	8.15%	5.77%
UBE3Adel 5	99.96%	94.91%	90.62%	4.29%	5.09%
UBE3Adel 6	99.92%	94.29%	89.80%	4.49%	5.71%
ASO1	99.88%	95.52%	91.62%	3.91%	4.48%
ASO2	99.69%	96.16%	91.76%	4.40%	3.84%
ASO3	99.91%	97.22%	92.65%	4.56%	2.78%
ASO4	99.58%	95.35%	91.23%	4.12%	4.65%
ASO5	99.44%	94.29%	89.40%	4.89%	5.71%
ASO6	99.31%	97.03%	92.14%	4.90%	2.97%
Scram1	99.80%	90.06%	86.08%	3.98%	9.94%
Scram2	99.89%	96.21%	91.68%	4.53%	3.79%
Scram3	99.79%	95.72%	91.61%	4.12%	4.28%
Scram4	99.55%	97.27%	92.98%	4.29%	2.73%
Scram5	99.80%	95.54%	91.47%	4.09%	4.46%
Scram6	99.51%	96.69%	92.38%	4.30%	3.31%

Table S4 mRNAseq analysis software versions

Software/Program	Version
FastQC	v0.11.5
Trimmomatic[104]	v0.36
Python	v2.7.8
MultiQC[105]	v1.1
STAR[106]	v2.5.3a
RStudio	v1.1.423
R	v3.5.0
DEseq2[107]	v1.20.0
org.Hs.eg.db	v3.6.0
AnnotationDbi	v1.42.1
ggplot2	v3.0.0
pheatmap	v1.0.10
RColorBrewer	v1.1-2
sva[98]	v3.28.0
vidger	v1.0.0
rgl	v0.99.16
rmarkdown	v1.10
dendextend	v1.8.0
gplots	v3.0.1
Vennerable	v3.1.0.9000
RBGL	v1.56.0
graph	v1.58.0
RDAVIDWebService[108]	v1.18.0
rJava	v0.9-10

Chapter 4

Examining the abundance and localization of human UBE3A protein isoforms in stem cells and stem cell-derived neurons

Carissa L. Sirois[1], Judy E. Bloom[2], James J. Fink[3], Dea Gorka[1], Noelle D. Germain[1], Justin Cotney[1], Leslie M. Loew[2], Eric S. Levine[3], Stormy J. Chamberlain[1]

[1] Department of Genetics and Genome Sciences, UConn Health, Farmington, CT
[2] Richard D. Berlin Center for Cell Analysis and Modeling, UConn Health, Farmington, CT
[3] Department of Neuroscience, UConn Health, Farmington, CT

This chapter is a manuscript being prepared for submission

My contributions: Work in this chapter was performed by the author aside from whole-cell patch clamp recordings, confocal microscopy of immunostained cells, electron microscopy, and UBE3A IP/ChIPseq

Abstract

Loss of expression of *UBE3A*, a gene regulated by genomic imprinting, causes Angelman Syndrome, a rare neurodevelopmental disorder. The *UBE3A* encodes an E3 ubiquitin ligase with three known protein isoforms in humans. Studies of the mouse isoforms suggest that the human isoforms may have differences in localization and neuronal function, and recently published AS case studies also indicate that the isoforms may play a role in AS. Here we have used CRISPR/Cas9 to generate human embryonic stem cells that lack the individual protein isoforms. We demonstrate that protein isoform 1 accounts for the majority of UBE3A protein in hESCs and neurons. We also show that all three isoforms predominantly localize to the cytoplasm in both cell types. Finally, we show that neurons lacking isoform 1 only recapitulate some of the phenotypes displayed by cultured AS neurons.

Introduction

Angelman Syndrome (AS) is a rare neurodevelopmental disorder that affects 1 in 15,000 individuals, characterized by severe seizures, intellectual disability, absent speech, ataxia, and happy affect.[20] The causative gene for AS, *UBE3A*, is regulated by tissue-specific genomic imprinting – it is expressed exclusively from the maternal allele in neurons and is expressed biallelically in other cell types.[11,21] The *UBE3A* gene encodes an E3 ubiquitin ligase that adds polyubiquitin chains to substrates, targeting them for degradation by the 26S proteasome.[25] In humans, there are three known UBE3A protein isoforms,[46] all of which are full length and therefore are presumably capable of functioning as an E3 ligase. Human isoforms 2 and 3 only differ from human isoform 1 by 23 and 20 amino acids at their N terminus, respectively.[48] While there has been little published research examining the human protein isoforms, studies performed in mouse indicate that the isoforms likely have differences in their localization and function.[49,50] It is also currently unknown whether the different isoforms play a role in AS. This is knowledge is important as some therapeutic avenues currently being explored for AS involve

delivery and expression of exogenous UBE3A transgenes using vector-based therapies.[71] For these approaches, knowledge of which protein isoforms need to be replaced in AS is essential.

Recently, Sadhwani and colleagues published three cases of AS caused by missense mutations at the isoform 1 translational start site.[77] Interestingly, these patients present with milder phenotypes than AS patients with 15q11-q13 deletions or other UBE3A loss of function mutations, including normal gait and use of syntactic speech. This data, coupled with the fact that the human and mouse isoforms are not entirely conserved, illustrates the need to study the human UBE3A protein isoforms specifically, as they may play a role in both normal neuronal function and in disease. Here, we have used human embryonic stem cells (hESCs) and their neuronal derivatives to examine the abundance and localization of the three human isoforms. We have also examined whether neurons lacking individual protein isoforms recapitulate any AS phenotypes previously seen in induced pluripotent stem cell (iPSC)- and hESC-derived neurons. We show that in both hESCs and hESC-derived neurons, all three UBE3A protein isoforms are predominantly localized to the cytoplasm, with low levels of expression in other cellular compartments. We also show that protein isoform 1 is the predominant protein isoform in human cells. Finally, we show that neurons lacking isoform 1 show some, but not all, of the phenotypes displayed by AS stem cell-derived neurons, while loss of isoform 2 or 3 does not produce any phenotypic changes. These results not only provide useful insight into the human UBE3A isoforms, but also provide information important for the development of potential AS therapies.

Methods

hESC culture and neural differentiation

hESCs were cultured on irradiated mouse embryonic fibroblasts and fed daily with hESC media (DMEM/F12 containing knockout serum replacement, L-glutamine + β-mercaptoethanol, non-essential amino acids, and basic fibroblast growth factor). hESCs were cultured in at 37°C in a humid incubator at 5% CO_2. Cells were manually passaged every 5-7 days.

hESCs were differentiated into neurons using a modified version of the monolayer protocol. Neural induction was begun 2 days after passaging by cultured cells in N2B27 medium (Neurobasal medium, 1% N2, 2% B27, 2 mM L-glutamine, 0.5% penicillin/streptomycin, 1% insulin-transferrin-selenium). N2B27 medium was supplemented with fresh Noggin (500 ng/mL) for the first 10 days of differentiation. Neural rosettes were manually passaged onto poly-D-lysine and laminin coated plates using the Stem Pro EZ passage tool approximately 14 days after beginning neural induction. Neural progenitors were replated at a high density around 3 weeks of differentiation, switched to neural differentiation medium (NDM) around 4 weeks of differentiation, then plated for terminal differentiation around 5 weeks. NDM consisted of neurobasal medium, 1% B27, 2 mM L-glutamine, 0.5% pen-strep, non-essential amino acids, 1 µM ascorbic acid, 200 µM cyclic AMP, 10 ng/mL brain-derived neurotrophic factor, and 10 ng/mL glial-derived neurotrophic factor. Neurons were maintained in culture until 10 to 17 weeks of differentiation.

CRISPR/Cas9-mediated genome editing

sgRNAs targeting the isoform translational start sites were designed using MIT's CRISPR design tool (http://crispr.mit.edu) and cloned into the pX459v2.0 vector (Addgene 62988), as previously described.[95,96] sgRNAs used to generate the UBE3A KO hESC line used here as a control have been described previously (**Chapter 2**). Prior to electroporation or nucleofection,

hESCs were treated with ROCK inhibitor (Y-27632; Selleck Chemicals) for 24 hours. hESCs were then singlized using Accutase (Millipore) and electroporated using the Gene Pulser X Cell (BioRad) or nucleofected using the Amaxa 4D Nucleofector (Lonza). For generation of the Isoform1 KO line, hESCs were electroporated in PBS with 10 µg of CRISPR plasmid and 8 µl of single-stranded oligonucleotide (ssODN) template. For generation of the Isoform2 KO and Isoform3 KO lines, hESCs were nucleofected with 2.5 µg of each CRISPR using the P3 Primary Cell Kit L (Lonza). hESCs were then plated onto puromycin-resistant (DR4) irrMEFs at low density, supplemented with ROCK inhibitor and L755507 (5 µM, Xcessbio), which has been shown to improve efficiency of homology directed repair. 24 hours after plating, cells underwent selection for 48 hours with puromycin (0.5 – 1 µg/ ml). Puromycin resistant colonies were screened by conventional PCR followed by restriction digest 11-14 days after plating. Putative clones were plated onto regular irrMEFs and successful genome editing was confirmed by Sanger sequencing.

Subcellular Fractionation

Subcellular fractionation was performed using the Cell Fractionation Kit – Standard (abcam) according to the manufacturer's instructions with the following modifications: 1) Protease Inhibitor Cocktail III was added to 1x Buffer A at the beginning of the fractionation protocol at a 1:1000 dilution; 2) a whole cell lysate was collected following the first lysis step by collecting 1/6 of the volume of lysate - downstream volumes used in the protocol were adjusted accordingly; and 3) whole cell and nuclear fractions were sonicated at the end of the protocol to sheer DNA using the following settings: [3 seconds on, 3 seconds off] x 2 at 30% amplitude. All fractions were stored at -80°C until use.

Western Blot

Equal volumes of lysate from each fraction were separated by SDS-PAGE using 4-20% TGX Stain-Free mini gels (BioRad). Protein was transferred to PVDF membrane using the TransBlot Turbo system (BioRad). Membranes were blocked in TBS Odyssey Blocking Buffer (LI-COR, Inc.) for 1 hour at room temperature then incubated in blocking buffer containing primary antibodies overnight at 4°C. Membranes were washed with TBS-T (Tris-buffered saline plus 0.1% Tween-20) at room temperature, incubated in blocking buffer containing IRDye Secondary Antibodies (LI-COR, Inc.) for 1 hour at room temperature, then washed again in TBS-T. All washes were done three times for 10-12 minutes each at room temperature. Membranes were imaged using the Odyssey imaging machine and software (LI-COR, Inc.). Images were quantified using Image Studio Lite (LI-COR, Inc.). The following primary antibodies were used: rabbit anti-UBE3A (1:3000; Bethyl A300-351), rabbit anti-UBE3A (1:3000; Bethyl A300-352), mouse anti-UBE3A (1:1000; Sigma E8655), rabbit anti-LMNB1 (1:5000; Abcam), mouse anti-GAPDH (1:10000; Millipore); rabbit anti-BAK (1:1000; Cell Signaling Technologies). The following secondary antibodies were used at a concentration of 1:10000: IRDye 800 CW Donkey anti-Rabbit, IRDye 680RD Goat anti-Rabbit, and IRDye 680RD Goat anti-Mouse (LI-COR, Inc.). A Western blot containing only nuclear fractions from all 4 hESC and neuron lines was performed with the following changes from above: blocking and antibody incubations were done in 5% BSA, anti-rabbit HRP-conjugated secondary antibody (1:3000; Cell Signaling Technologies) was used, and the blot was imaged using Clarity Western ECL substrate (BioRad) on the ChemiDoc Touch imaging system (BioRad).

Immunocytochemistry

hESCs and neurons were grown on glass chamber slides for immunocytochemistry. Cells were fixed using room temperature 4% paraformaldehyde for 10 min then permeabilized using 0.5% PBS-Triton X 100 (PBS-T) for 10 min at room temperature. Cells were blocked in 0.1% PBS-T containing 2% bovine serum albumin and 5% normal goat serum. Cells were incubated in

primary antibody in blocking buffer overnight at room temperature then washed with PBS. Cells were then incubated in secondary antibody in blocking buffer for 3 hours at room temperature then washed with PBS. Cells were mounted with ProLong Gold Anti-Fade Hard Set with DAPI (ThermoFisher Scientific) and allowed to set overnight at room temperature before imaging. The following primary antibodies were used: mouse anti-UBE3A (1:500; Sigma E8655), rabbit anti-OCT4 (1:200; Stemgent 09-0023), chicken anti-MAP2 (1:10000; abcam), rabbit anti-NeuN (1:300; abcam). The following secondary antibodies were used: anti-mouse Alexa Fluor 488 (1:400; Invitrogen), anti-rabbit Alexa Fluor 594 (1:400; Invitrogen), anti-chicken Alexa Fluor 647 (1:250, abcam). Confocal images (134.7 x 134.7um) were acquired with a 63x oil immersion lens (numerical aperture 1.4) on a Zeiss 780 confocal system mounted on an inverted Axio Observer Z1 (Carl Zeiss, Germany). Pixel dwell time was 1.27 seconds, scan time of 2.34 seconds and images were averaged 2 times. Three randomly chosen areas of the chamber slide were images for each cell line. Z-series were done with a 0.2 µm interval and maximum projection images were created using Zeiss ZEN software.

Electrophysiology

Neurons were plated onto glass coverslips around 5 weeks of differentiation. Whole-cell voltage and current clamp recordings were performed at 12 weeks of differentiation as previously described.[76]

qRT-PCR

cDNA was synthesized using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) according to manufacturer's instructions. Quantitative RT-PCR was performed using Taqman Gene Expression Assays and Mastermix (Thermo Fisher Scientific) on the Step One Plus (Thermo Fisher Scientific). Reactions were performed in technical duplicates, with

GAPDH Endogenous Control Taqman Assay used as the housekeeping gene for normalization. Gene expression was quantified using the $\Delta\Delta C_t$ method.

Immunogold Transmission Electron Microscopy

Neurons were plated for terminal differentiation onto plastic 4 well plates at 6 weeks of differentiation. At 10 weeks, neurons were fixed using 4% paraformaldehyde at room temperature for 30 min. Cells were washed in PBS at room temperature, then blocked and permeabilized by incubating in 0.1% PBS-T containing 2% bovine serum albumin and 5% normal goat serum for 30 min at room temperature. Cells were incubated overnight at room temperature in blocking buffer containing primary antibody (mouse anti UBE3A (Sigma) at 1:250). Cells were washed at room temperature in PBS then incubated in secondary antibody for 1 hour at room temperature. Secondary antibody (goat anti-mouse Nanogold Fab) was diluted 1:200 in 1% BSA. Cells were washed several times in PBS then washed in deionized water. Gold enhancement was performed for 4 min using GoldEnhance (Nanoprobes). Cells were rinsed in deionized water to stop enhancement reaction.

To prepare cells for electron microscopy, cells were rinsed for 5 min in 0.1 M Cacodylate buffer, then post-fixed in 1% OsO_4, 0.8% Potassium Ferricyanide in 0.1 M Cacodylate buffer for 30 min at room temperature. Cells were then washed 5 times in deionized water for 5 min per wash, block stained in 1% Uranyl Acetate in deionized water for 30 min at room temperature, then washed again 3 times (5 min each) in deionized water. Next cells were dehydrated in ethanol (5 min in 50% EtOH, 5 min in 75% EtOH, 5 min in 95% EtOH, then 3 incubations in 100% EtOH for 5 min each), then infiltrated in 100% resin overnight at room temperature. The following day, fresh 100% resin was added and the cells were polymerized at 60°C for 48 hours. Thin sections of 70-80 nm were cut on the Ultramicrotome Leica EM UC7 and sections placed on Cu grids. Sections were counterstained with 6% Uranyl Acetate in 50% methanol for 4 min. Images were acquired with the Hitachi H-7650 Transmission Electron Microscope. Electron microscopy

preparation and imaging was performed by the UConn Health Central Electron Microscopy Facility.

ChIPseq

ChIPseq was performed as previously described [109,110] on hESCs, hESC-derived neural progenitors (NPCs; 5 weeks of differentiation), and hESC-derived neurons (10 weeks of differentiation). Antibodies used in this experiment were as follows: rabbit anti-UBE3A (Bethyl A300-351), rabbit anti-UBE3A (Bethyl A300-352), anti-H3K27ac (ab4729, Abcam). Two replicate IPs were performed per antibody using approximately 100 µg of chromatin and 10 µg of antibody per IP.

Results

Generation of isogenic hESC lines lacking individual protein isoforms

To study the abundance and localization of the UBE3A protein isoforms, we first generated isogenic hESC lines lacking the individual protein isoforms. Isogenic hESCs and neurons would allow us to minimize molecular and phenotypic differences caused by normal human genetic variation. All three human protein isoforms are full length versions of UBE3A: isoforms 2 and 3 have an additional 23 and 20 amino acids, respectively, at their N terminus. Because of this, the RNAs encoding the isoforms are nearly identical aside from the exons that encode for these additional amino acids.[48] The three UBE3A protein isoforms are, however, translated from unique translational start sites.[47] We took advantage of these three sites in order to generate three isogenic hESC lines, each of which lacked an individual protein isoform. Using the CRISPR/Cas9 system, we mutated one of the three translational start sites, changing it from a methionine to a leucine, which would prevent the translation of the isoform of interest (**Figure 1**). Because the methionine that serves as the isoform 1 translational start site is present in the other two protein isoforms, we used multiple protein prediction softwares[86,87] to find an amino acid substitute (leucine) that was predicted to have a benign effect on the other two isoforms (**Supplemental Figure 1**).

Isoform 1 is the most abundant UBE3A protein isoform

To determine the relative abundance of the human UBE3A protein isoforms in hESCs and hESC-derived neurons, we examined total UBE3A protein levels in whole cell lysates prepared from the isoform knockout and isogenic control lines. Our data indicate that isoform 1 is the predominant isoform in both cell types – loss of this isoform produced a significant reduction in total UBE3A levels in whole cell lysates prepared from both hESCs (**Figure 2A/B**) and neurons (**Figure 3A/B**). Loss of isoform 2 or isoform 3, however, did not produce any significant changes in total UBE3A levels. These data indicate that isoform 1 is the most abundant of the three human protein isoforms in hESCs and hESC-derived neurons.

All three UBE3A isoforms are predominantly localized to the cytoplasm

Because studies of the mouse protein isoforms indicate differences in isoform localization,[49] we sought to determine whether the human protein isoforms also localized differentially. Using a subcellular fractionation protocol, we examined UBE3A abundance in cytoplasmic, mitochondrial, and nuclear fractions in both hESCs and neurons. In both cell types, UBE3A appears to localize predominantly to the cytoplasm, both in normal and isoform-null lines (**Figure 2, Figure 3**). Consistent with our whole cell results, loss of isoform 1 produced significant loss of UBE3A in the cytoplasmic fraction (**Figure 2C, Figure 3C**), but not the mitochondrial or nuclear fractions (**Figure 2D/E, Figure 3 D/E**). Loss of isoform 2 or isoform 3 did not cause a significant change in the abundance of UBE3A protein in any cellular fractions when compared to the H9 parent line (**Figure 2A-E, Figure 3A-E**).

To confirm these results, we also examined UBE3A localization via immunocytochemistry. We first stained hESCs for UBE3A and OCT3/4, a transcription factor and pluripotency marker known to localize to the nucleus (**Figure 2F, Supplemental Figure 3**). Here we included a UBE3A KO hESC line to serve as an additional control (**Supplemental Figure 3**). Consistent with our fractionation results, expression of UBE3A in hESCs appears to be predominantly cytoplasmic, as there is expression of UBE3A outside of the area positive for OCT3/4 staining. We next stained hESC-derived neurons for UBE3A and MAP2, a microtubule protein that localizes to the cytoplasm and is a marker for post-mitotic neurons (**Figure 3C, Supplemental Figure 4A**). Although the strongest UBE3A signal appears in the nucleus, there is still obvious UBE3A signal in the processes and in the area of the soma outside of the nucleus (**Supplemental Figure 4A & 5**).

Studies of UBE3A localization in mouse brain indicate that UBE3A is initially expressed in the cytoplasm but becomes increasingly nuclear upon postnatal maturation.[111] To determine whether the intense staining in the soma was nuclear UBE3A, we stained neurons for UBE3A and NeuN, which is predominantly expressed in the nucleus and, to a lesser extent, in the

processes, of mature neurons. Comparison of the UBE3A signal with both NeuN and DAPI shows that the intense UBE3A signal in the soma colocalizes with both of these nuclear markers (**Supplemental Figure 4B**), but only to a certain extent (**Supplemental Figure 4B, arrows**).

Because the results of the immunostaining show a stronger UBE3A nuclear signal than the subcellular fractionations, we sought to further confirm the results from each procedure. We first confirmed the results of our fractionation protocol by using a second UBE3A primary antibody, which reacts to a different portion of the protein (N terminus versus HECT domain; **Supplemental Figure 6A**). Next, we wanted to see if perhaps our nuclear fractions were simply too diluted or run at too small of a volume to see robust nuclear UBE3A bands. To do this, we ran a Western blot using a larger volume of lysate that was intentionally exposed to the point of saturation (3 min; **Supplemental Figure 6B**), which showed a faint signal. Finally, we performed a Western blot using the same primary antibody that is used for the immunostaining (**Supplemental Figure 6C**). Interestingly, although we do not see a band in the nuclear fractions at the expected UBE3A size of 100 kDa, we do see a band that only appears in the whole cell and nuclear fraction lysates that is less than 75 kDa in size. It is possible that this unknown 75 kDa protein is also causing the increased nuclear signal when we perform immunostaining, as this band does not appear when using other UBE3A antibodies (for comparison, see **Figure 2A, 3A, and Supplemental Figure 6A**).

To confirm the results of our immunostaining, we examined UBE3A localization in neurons by transmission electron microscopy. These results showed diffuse UBE3A puncta in the soma both within and outside of nucleus (**Supplemental Figure 7A-B**), as well as in the neuronal processes (**Supplemental Figure 7C**). Finally, we investigated whether the small proportion of UBE3A that is localized to the nuclear is capable of acting as a transcriptional co-activator.[40] ChIP-seq for UBE3A using two different primary antibodies showed that UBE3A did not preferentially bind to DNA in NPCs or neurons (**Supplemental Figure 8A**).

Isoform 1-null neurons recapitulate some AS phenotypes

We last wanted to determine whether neurons lacking the individual UBE3A protein isoforms displayed any *in vitro* AS phenotypes. Based on our results examining the abundance of the isoforms (**Figure 3A-B**), and our knowledge of the AS patients with isoform 1 translational start site mutations,[77] we hypothesized that loss of isoform 1 would produce AS phenotypes in hESC-derived neurons while loss of the other two isoforms would likely would not produce any phenotype. Previously we established that AS iPSC-derived neurons exhibit a phenotype of impaired electrical maturation, as indicated by a more depolarized resting membrane potential (RMP), immature patterns of action potential firing, and decreased frequency of spontaneous excitatory synaptic activity.[76] We examined these three phenotypes in our isoform-null and isogenic control neurons at 12 weeks of differentiation. Isoform 1 KO neurons displayed a more depolarized resting membrane potential compared to controls, while there were no differences in the isoform 2 KO or isoform 3 KO neurons (**Figure 4A**). Interestingly, there were no differences in action potential firing (**Figure 4B**) or synaptic activity (**Figure 4C**) in the any of the isoform-null neurons, indicating that the isoform 1 KO neurons only display some AS electrical phenotypes.

Finally, we wanted to examine whether the isoform-null neurons displayed any molecular AS phenotypes. Recently we established and validated a list of genes that are consistently differentially expressed in AS stem cell-derived neurons when compared to isogenic controls by mRNAseq. We chose to examine the expression of a handful of these AS genes in the isoform 1 KO neurons, as they most closely resemble AS neurons in terms of their UBE3A protein levels and electrical phenotype. We examined the expression of 6 genes: *MGP, GJA5, TBX18, TBR1, NEUROD2,* and *EDNRA* (**Figure 4D**). Four of these genes (*TBX18, TBR1, NEUROD2, EDNRA*) were significantly differentially expressed in the Isoform 1 KO neurons in the same manner as AS neurons (Student's t-test, $p < 0.05$), while *GJA5* was not different between the two lines.

Unexpectedly, *MGP* was actually decreased in the Isoform 1 KO neurons. Together these data indicate that the Isoform 1 KO neurons recapitulate some, but not all, AS phenotypes *in vitro*.

Discussion

The human *UBE3A* gene encodes for an E3 ubiquitin ligase that has three known protein isoforms.[46] Surprisingly little is known about the human isoforms, however studies in mouse indicate that the isoforms may have differences in localization and function.[49,50] Here we have used the CRISPR/Cas9 system to generate hESCs lacking these individual isoforms and studied their abundance and localization in both hESCs and neurons. We have also examined the phenotype of neurons lacking the individual isoforms. We have shown that human isoform 1 is the most abundant of the three isoforms, accounting for almost ninety percent of total UBE3A protein in these two cell types. We have also shown that in both hESCs and neurons, UBE3A localizes predominantly to the cytoplasm, independent of which isoform is absent or present, with low levels of expression in the nucleus. Finally, we have shown that neurons lacking isoform 1 display some *in vitro* AS phenotypes, in terms of their electrical properties and gene expression.

To examine the localization of the UBE3A isoforms, we used two approaches: subcellular fractionation and immunocytochemistry. Both approaches showed that isoform 1 is the predominant protein isoform and that all isoforms are most abundant in the cytoplasm. Loss of individual isoforms did not change the localization of UBE3A using either approach. Interestingly, we appeared to see more UBE3A in the nucleus of normal and isoform-null neurons via immunostaining, as shown by colocalization with DAPI and NeuN, than would be anticipated based on the results of our fractionations. However, the amount of UBE3A that we see in normal neurons by TEM appears to agree with the fractionation results.

We believe that the strong UBE3A nuclear UBE3A signal that we see by immunostaining is likely due to the fact that the small amount of UBE3A that is localized to the nucleus is contained within a much smaller volume than the cytosolic UBE3A, which is spread out across

the volume of the entire neuron and its processes. It is also possible, based on our Western blot using the same primary antibody, that the strong nuclear UBE3A signal is related to the >75 kDa band that we see via Western blot. Based on our knowledge of UBE3A localization in the developing mouse brain,[111] it is also possible that human UBE3A becomes increasingly nuclear during the course of human brain development, but that our neurons are developmentally too immature to display these changes in localization.

Interestingly, although isoform 1-null neurons lack the majority of the total UBE3A protein, they do not fully recapitulate AS electrical and molecular phenotypes. Isoform 1-null neurons have a more depolarized resting membrane potential compared to the other three neuron lines, however, the RMP of these neurons would still be considered mature for their stage of *in vitro* differentiation. We hypothesize that the difference in RMP between isoform 1-null neurons and their isogenic controls would be even more pronounced at an earlier age, such as 9 to 10 weeks *in vitro*. Additionally, this more mature RMP seen in all four neuron lines is also consistent with their action potential firing pattern, as most neurons were capable of firing single or mature trains of action potentials. Another possibility is that, even though loss of isoform 1 dramatically reduces total UBE3A levels in our stem cells and neurons, *in vitro* AS phenotypes are truly the result of loss of two or more of the protein isoforms. Future studies should examine the effects of the combined loss of two isoforms at a time to further elucidate their role in the electrical maturity of normal and AS neurons, as well as examine these phenotypes at various stages of neuronal development in our isoform-null neurons.

We have also shown here that isoform 1 accounts for the majority of UBE3A protein in hESCs and neurons, while isoforms 2 and 3 appear to account for very little protein at this stage. This knowledge of the isoforms' abundance is important for the development of AS therapies. One promising therapeutic avenue currently being explored for AS is the introduction of a UBE3A transgene through vector-based therapies,[71] which typically can only contain one cDNA at a time. Replacement of all three human isoforms using this approach, therefore, would

likely require the development of three separate vectors. Based on the results shown here, it is possible that delivery of isoform 1 alone may be a useful therapeutic approach, as it accounts for the majority of total UBE3A protein in human neurons *in vitro*. Future studies should examine whether introduction of individual UBE3A protein isoforms into human AS neurons *in vitro* is capable of restoring their phenotypes.

In summary, we have used isogenic isoform-null hESCS and hESC-derived neurons to examine the relative abundance and localization of the human UBE3A protein isoforms. We have shown that isoform 1 is the most abundant, and that most UBE3A protein localizes to the cytoplasm in both cell types, independent of the specific isoform. We have also demonstrated that neurons lacking isoform 1 recapitulate some, but not all, *in vitro* AS phenotypes. This knowledge is not only important for the study of AS, but also the development of potential AS therapies.

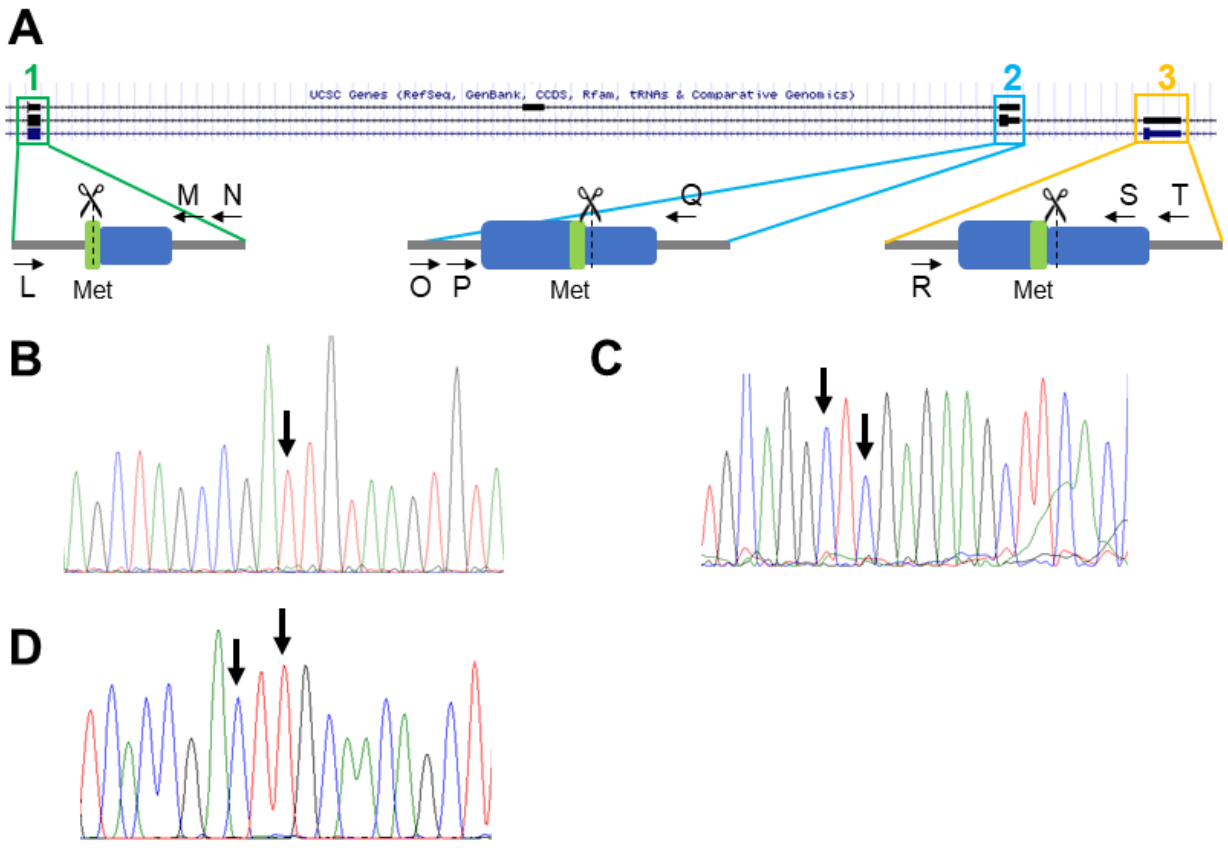


Figure 1. Generation of isogenic isoform-null hESC lines

A Schematic illustrating isoform translational start sites and proposed genome editing. Arrows indicate primers used for screening. Scissors indicate CRISPR cut sites. Blue box = exon; grey line = intron; green = methionine used as start site **B-D** Sanger sequencing showing mutation of translational start site (ATG) for isoform 1 (**B**), isoform 2 (**C**), and isoform 3 (**D**). Arrows indicate locations of changed nucleotides.

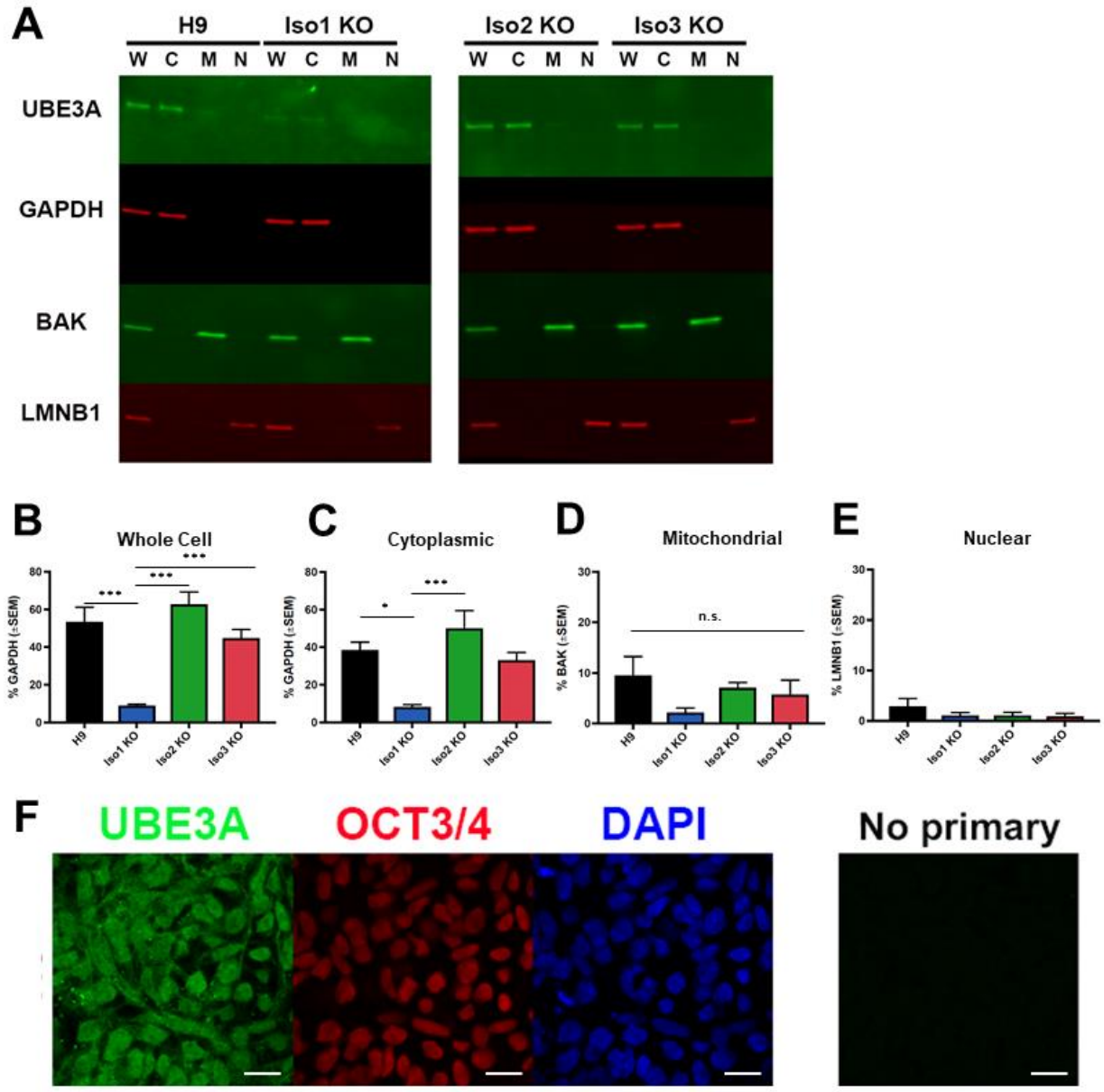


Figure 2. Abundance and localization of UBE3A isoforms in hESCs
A Western blot showing UBE3A levels in subcellular fractions from H9 and isoform-null hESC lines. W = whole cell lysate, C = cytoplasmic fraction, M = mitochondrial fraction, N = nuclear fraction. **B-E** Quantification of Western blots for each fraction. Graphs show average percentage of appropriate loading control (n = 4 fractionations). Error bars: standard error of the mean. * $p < 0.05$ *** $p < 0.005$ (univariate ANOVA) **F** Immunocytochemistry for UBE3A in H9 hESCS shows that UBE3A is cytoplasmic, in agreement with above fractionation results. Scale bar 20 µm

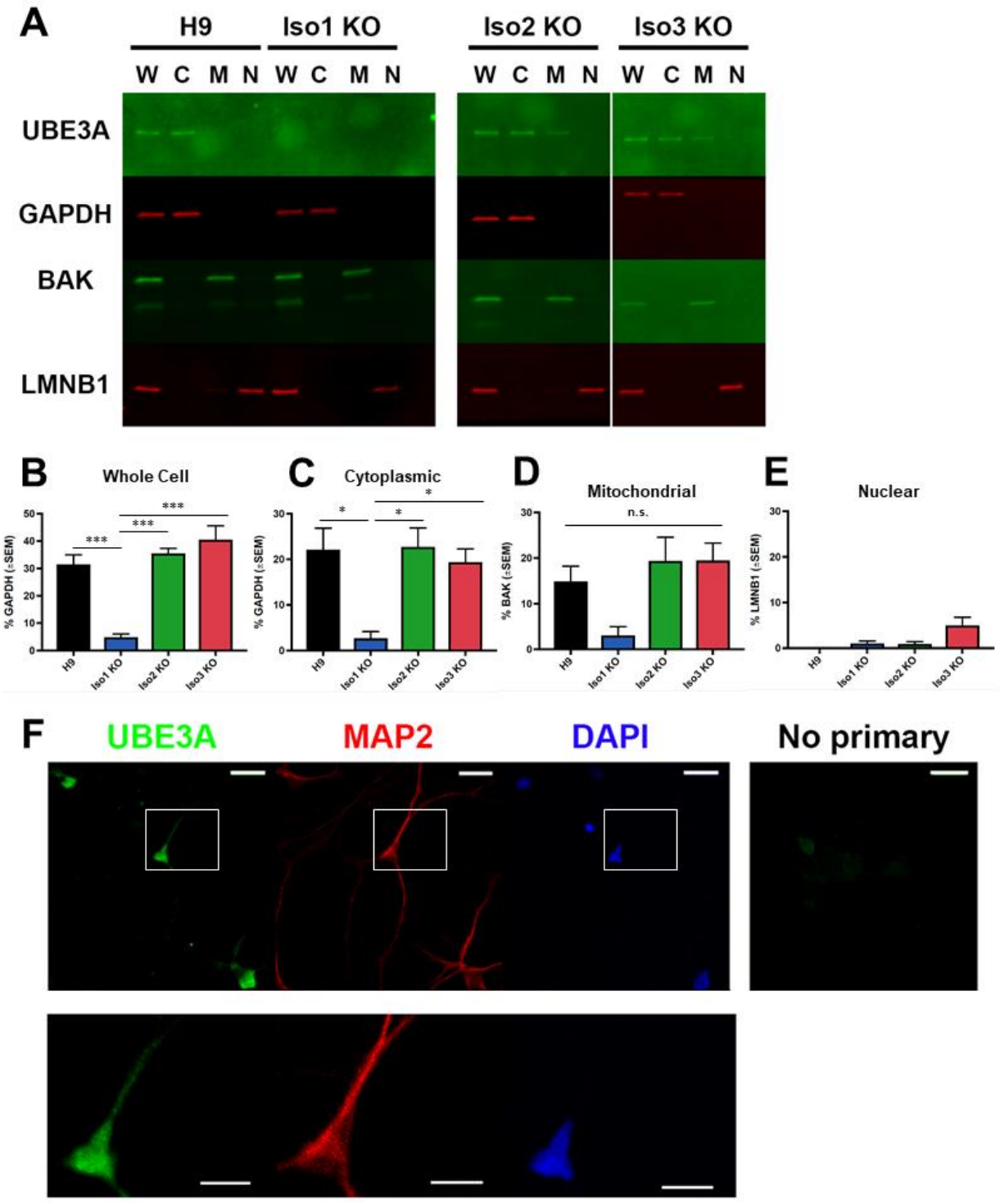


Figure 3. Abundance and localization of UBE3A isoforms in neurons
A Western blot showing UBE3A levels in subcellular fractions from H9 and isoform-null neurons. W = whole cell lysate, C = cytoplasmic fraction, M = mitochondrial fraction, N = nuclear fraction. **B-E** Quantification of Western blots for each fraction. Graphs show average percentage of appropriate loading control (n = 3 fractionations). Error bars: standard error of the mean.

* $p < 0.05$ *** $p < 0.005$ (univariate ANOVA) **F** Immunocytochemistry for UBE3A in H9 neurons shows that UBE3A is highly concentrated in the nucleus (top), but also is expressed in the cytoplasm (bottom). Box in top row indicates region in bottom row. Scale bar 20 μm (top), 10 μm (bottom).

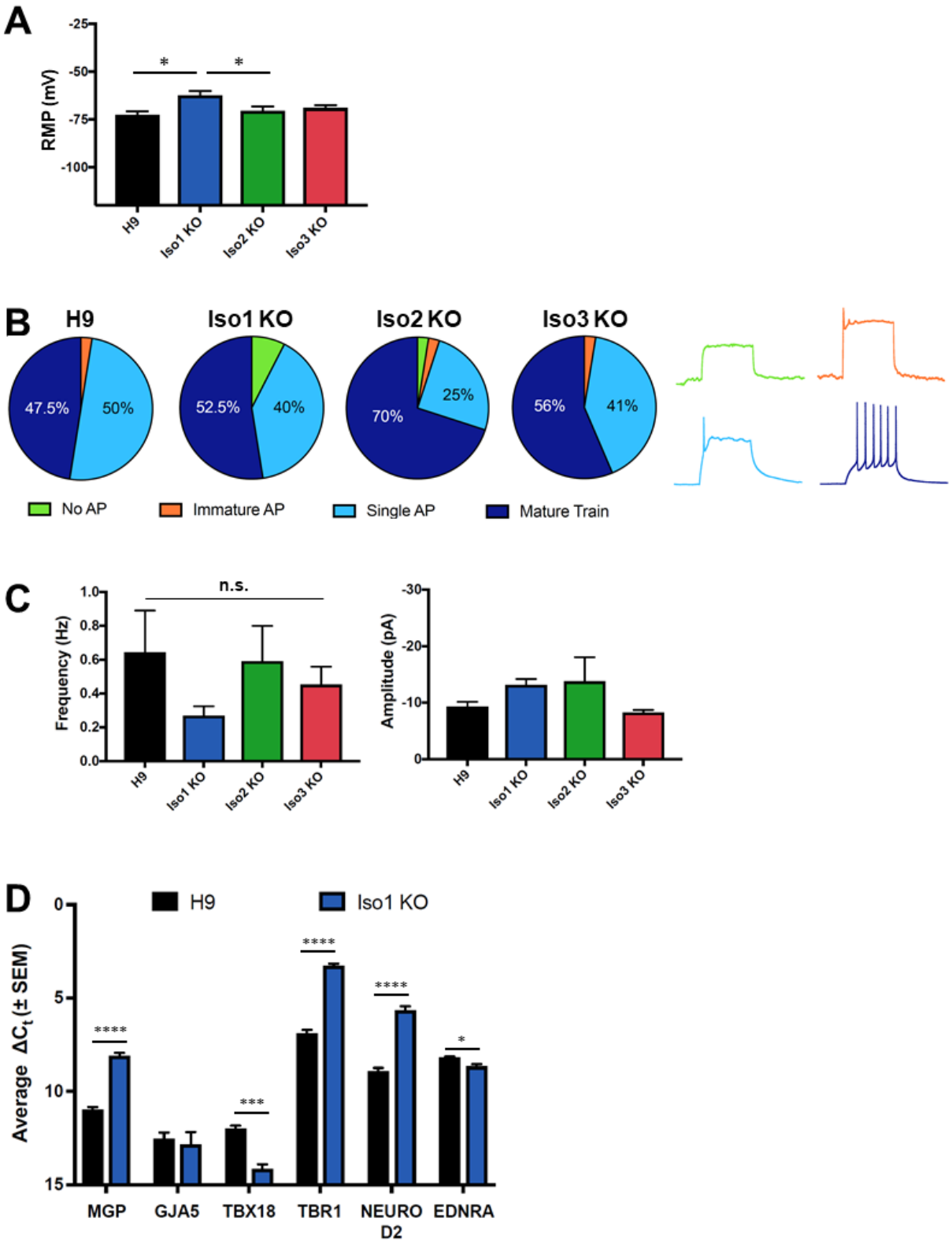


Figure 4. Isoform 1-null neurons partially recapitulate AS cellular phenotypes
A Isoform-1 null neurons have a more depolarized resting membrane potential. * $p < 0.05$ (univariate ANOVA) **B** No significant differences in … (continued on next page)

action potential (AP) firing upon loss of any UBE3A isoforms Left: AP characterization in isoform-null neurons; right: sample AP traces **C** Isoform-null neurons do not have any significant differences in spontaneous excitatory synaptic activity. n = 39-40 cells on 3 coverslips per cell line. **D** Loss of isoform 1 affects expression of some, but not all, genes differentially expressed in AS neurons, as assayed by qRT-PCR * $p < 0.05$, *** $p < 0.005$, **** $p < 0.001$ (student's t-test)

Figure S1. Effect of proposed amino acid change on isoform 2 (A) and isoform 3 (B) protein. Top: results from Predict Protein software showing effects of every possible amino acid substitution of the methionine. Bottom: results from Predict SNP software showing predicted effects of L to M substitution from 7 different structure prediction algorithms.

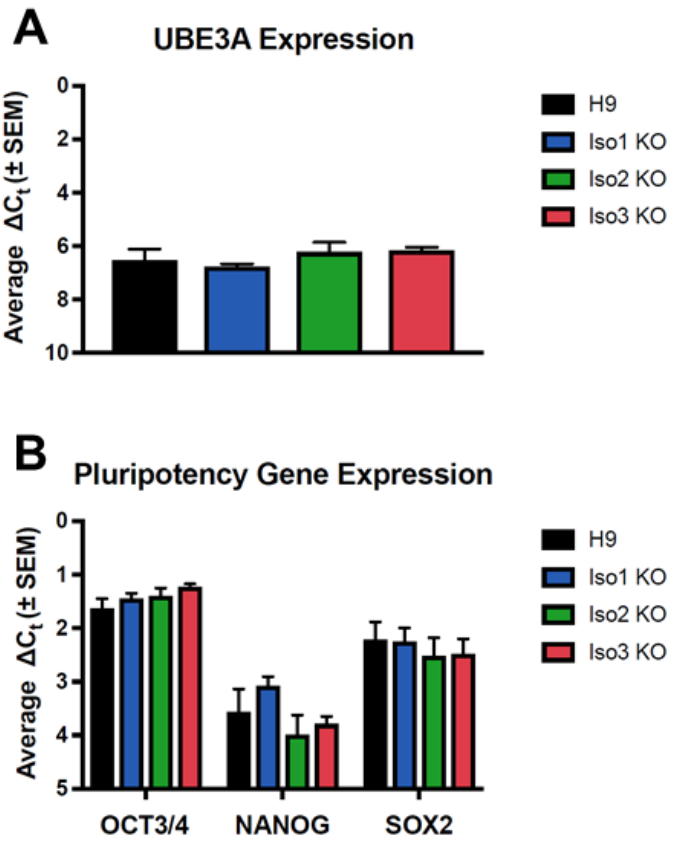


Figure S2. Characterization of isoform-null hESCs
A qRT-PCR results showing UBE3A expression in H9 and isoform-null hESCs **B** qRT-PCR results showing expression of pluripotency genes in H9 and isoform-null hESCs

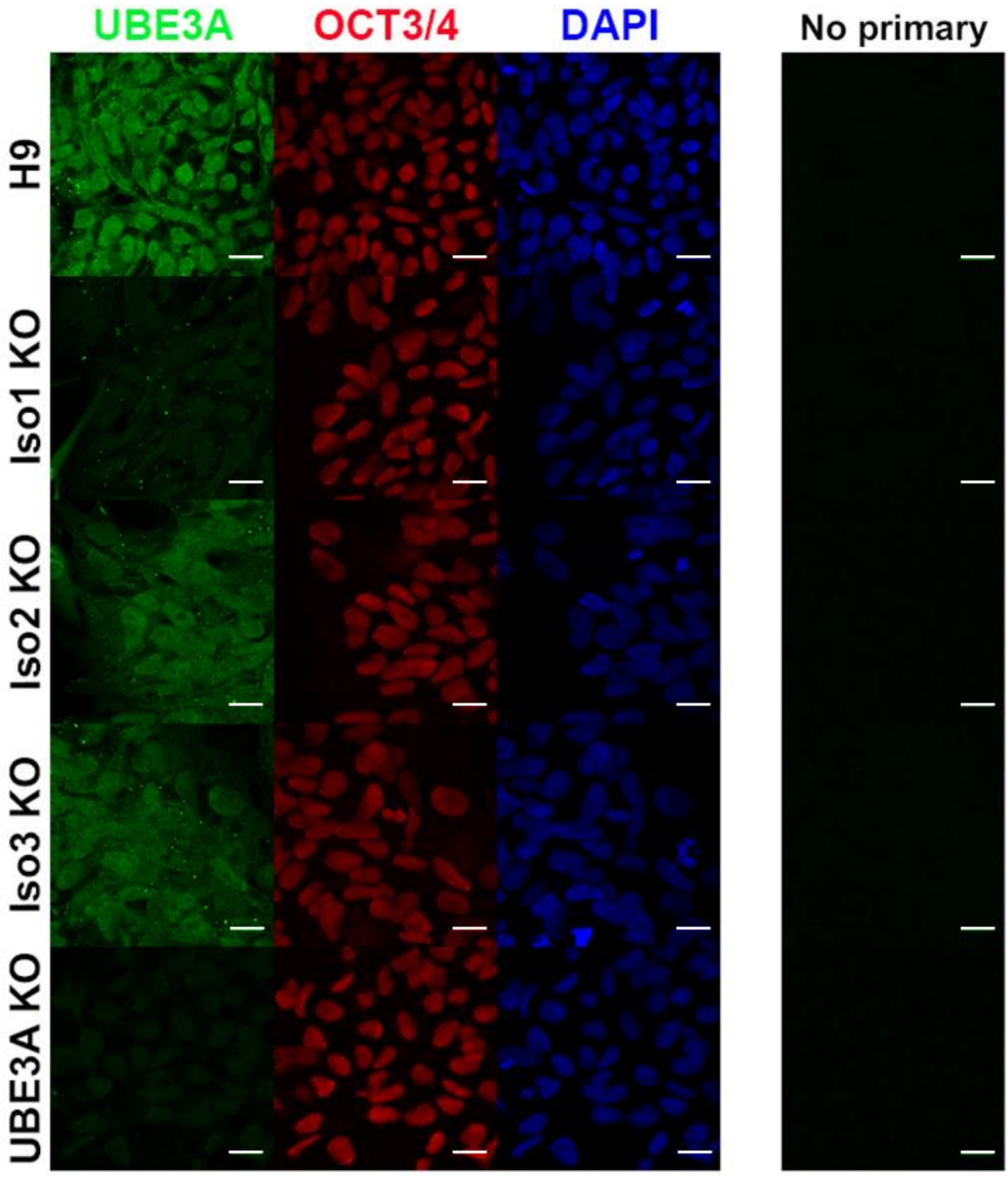


Figure S3. UBE3A expression in isoform-null hESCs by immunocytochemistry

Abundant expression of UBE3A is seen in the cytoplasm in all isoform-null cell lines, indicating that the localization of UBE3A does not differ upon loss of any isoform. Isoform 1-null hESCs show dramatically reduced protein levels. UBE3A KO hESCs were included as a negative control. OCT3/4 staining was included to show localization of a known nuclear protein. Scale bar 20 μm

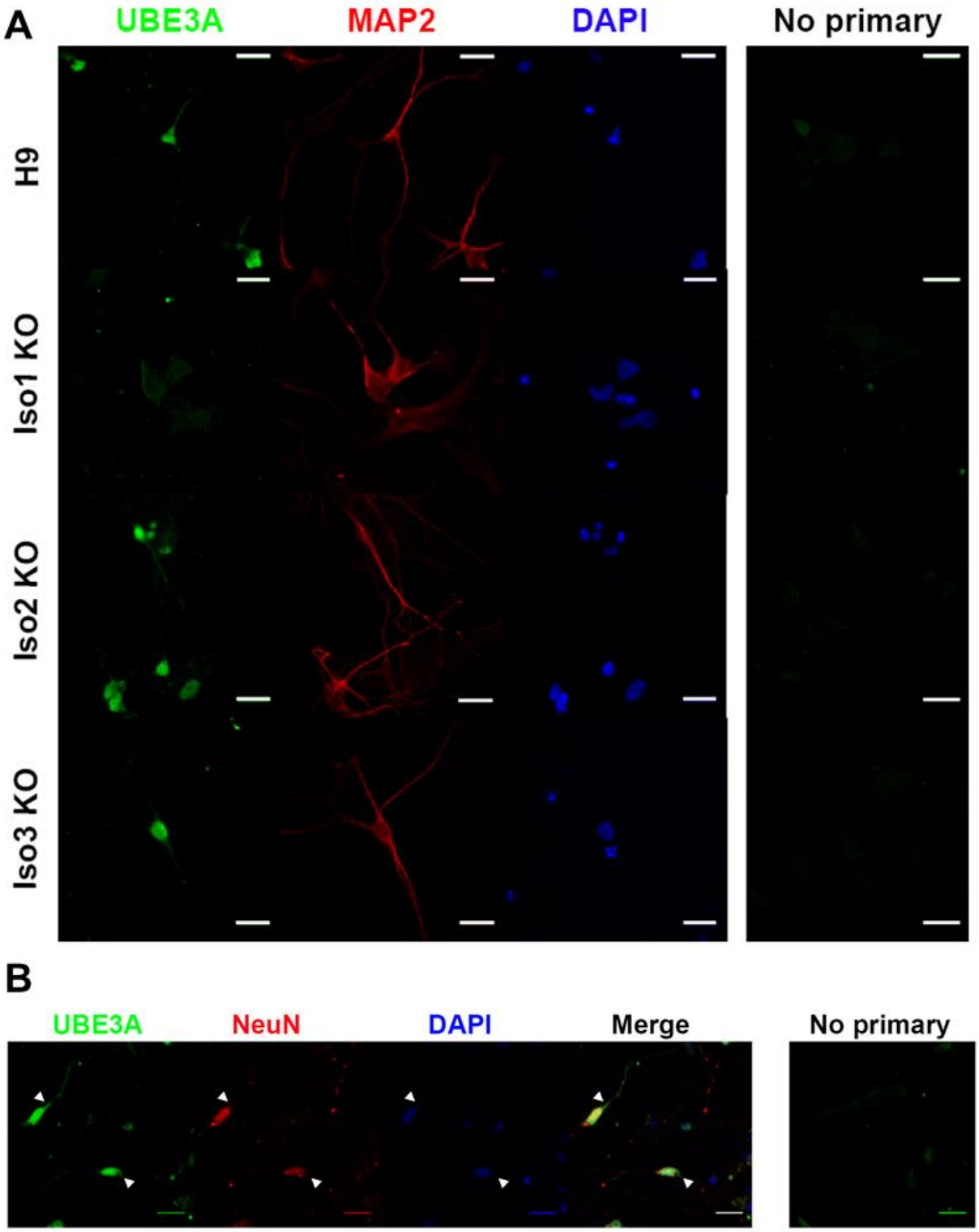


Figure S4. UBE3A expression in neurons by immunocytochemistry
A UBE3A signal appears to localize predominantly to the nucleus by immunostaining in neurons, in contradiction to fractionation results. **B** Comparing localization of UBE3A with localization of NeuN, a nuclear marker, shows strong UBE3A signal outside of the nucleus (arrows). Scale bar 20 μm

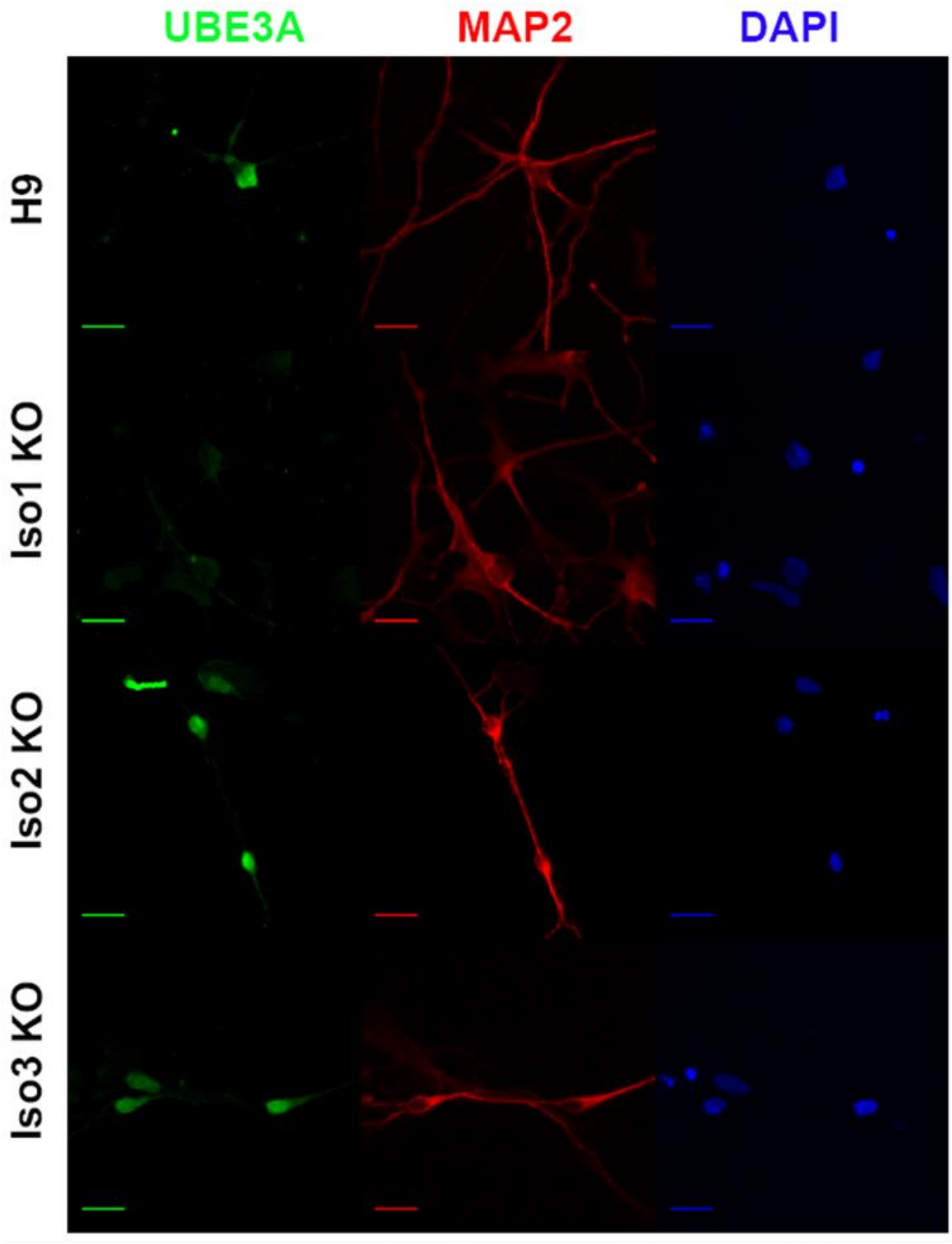


Figure S5. Maximum projection images from Z stack shows cytoplasmic UBE3A expression in all four neuron lines Scale bar 20 μm

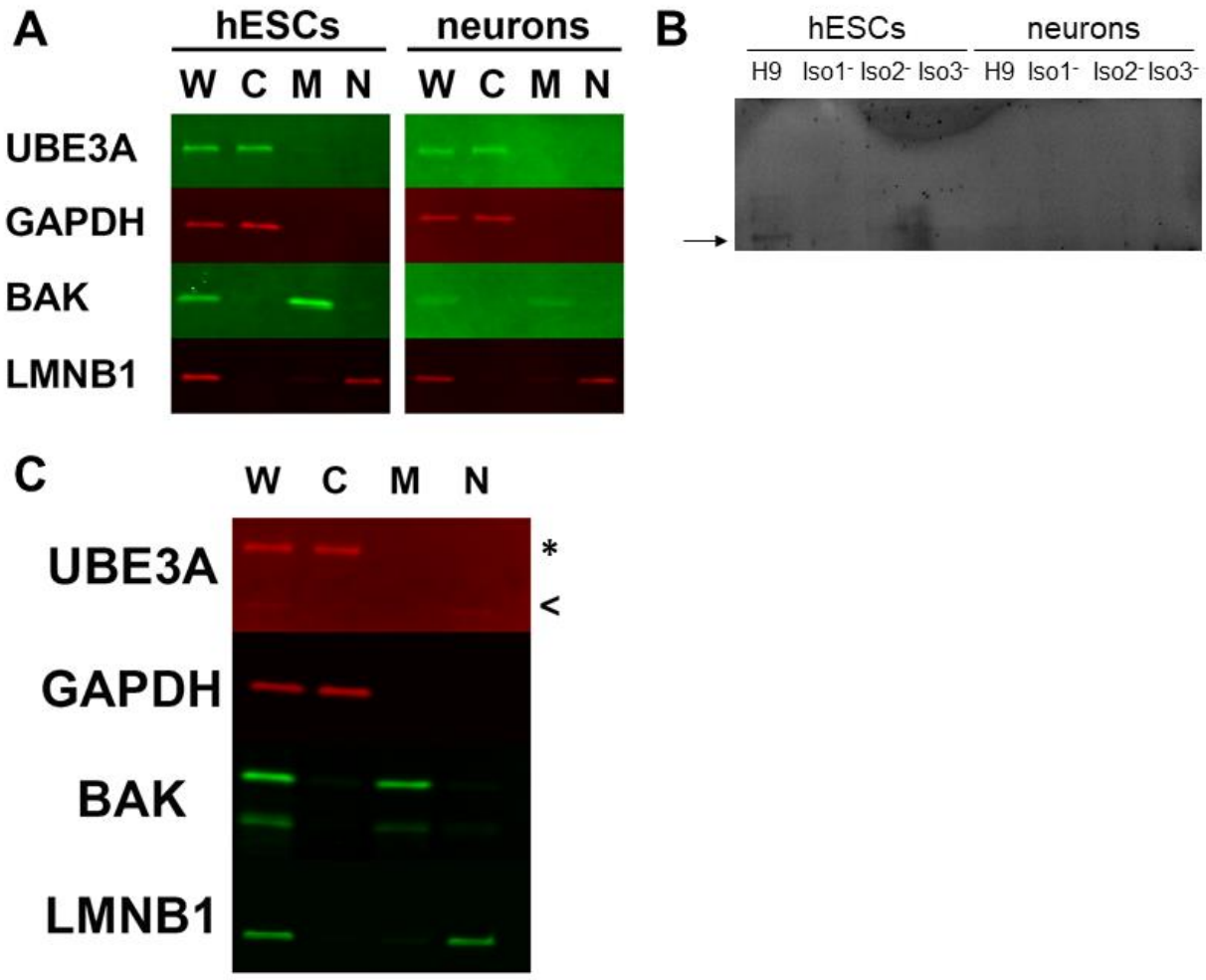


Figure S6. Additional UBE3A Western blots
A Confirmation of fractionation results in H9 hESCs and neurons using a secondary UBE3A primary antibody **B** Overexposure of nuclear fractions from hESCs and neurons show faint nuclear UBE3A signal **C** Western blot using same primary antibody used for ICC on lysates from H9 neurons confirms findings from fractionations. Presence of smaller nuclear band could explain intense nuclear signal in ICC images. * expected UBE3A size, < shorter band seen only in whole cell and nuclear fractions

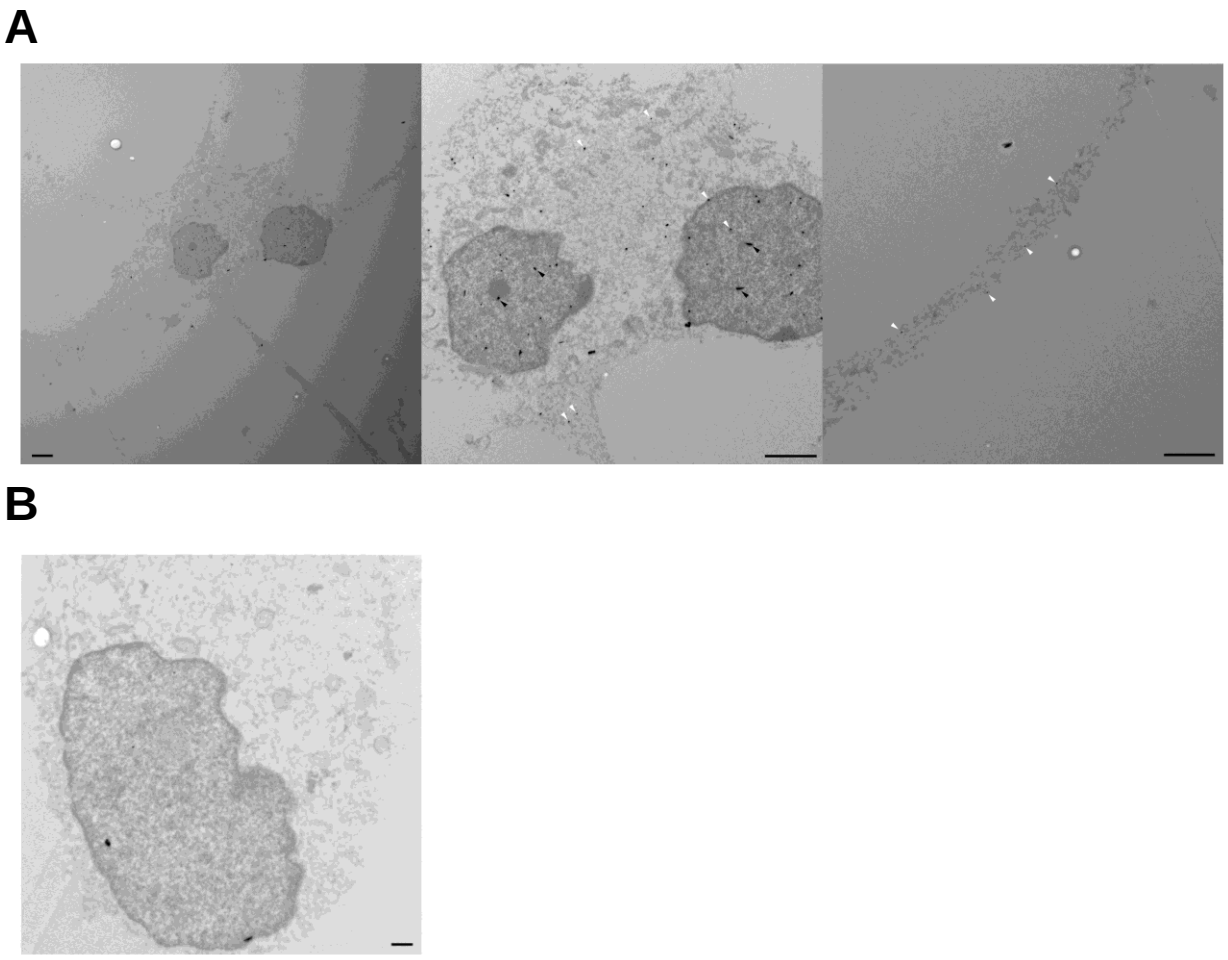


Figure S7. Transmission electron microscopy images show diffuse UBE3A localization in cytoplasm and nucleus in H9 neurons
A Immunogold labeling shows diffuse UBE3A staining (white triangles) in the nucleus and soma (middle image) as well as in the neuronal processes (right image). Examples of non-specific puncta are indicated by black arrows. Scale bar 2 μm **B** No primary control image. Scale bar 500 nm

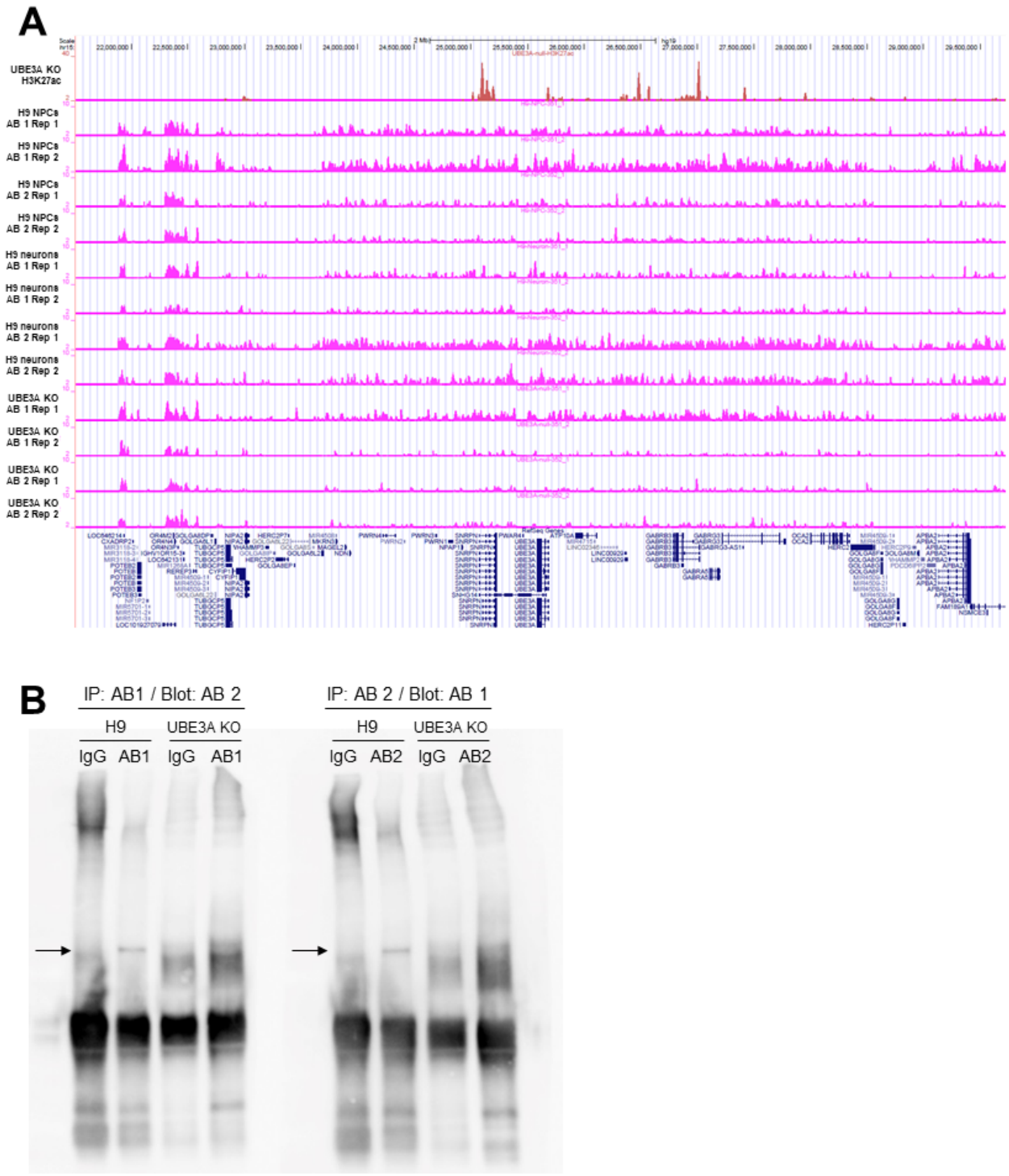


Figure S8. ChIPseq indicates that nuclear UBE3A does not directly bind to DNA in NPCs or neurons

A ChIPseq peaks in 15q11-q13 region. Top (red): H3K27ac ChIPseq performed in UBE3A KO hESCs as a positive control. Bottom (pink): UBE3A ChIPseq indicates no specific binding anywhere in the genome (15q11-q13.1 region shown here as an example) in either H9 NPCs or H9 neurons. UBE3A KO hESCs were included as a negative control. **B** IP showing that the antibodies used for ChIPseq are capable of … (continued on next page)

pulling down UBE3A protein. Left image: IP was performed with antibody 1 then the membrane was blotted with antibody 2. Right image: IP was performed with antibody 2 then the membrane was blotted with antibody 1.

Table 1. sgRNAs used to generate stem cell lines	
Cell Line	sgRNA sequence + **PAM**
Isoform 1 start site mut.	**CCG**AATGTAAGTGTAACTTGGTT
Isoform 2 start site mut.	AAAAGGAGTGGCTTGCAGGA**TGG**
Isoform 3 start site mut.	ATCACCCTGATGTCACCGAA**TGG**

Table 2. ssODNs used to generate stem cell lines	
Cell Line/Project	ssODN sequence
Isoform 1 start site mutation (H9)	AGAACCTCAGTCTGACGACATTGAAGCTAGTCGATTGTAA GTGTAACTTGGTTGAGACTGTGGTTCTTAT
Isoform 2 start site mutation (H9)	ATTCAAATGGTGGCTCACTTCCAATAACACTGGTGAAGCTTCTCGA GCCTGCAAGCCACTCCTTTTACCTCCACTGTAACTCTCTAGGAGAG
Isoform 3 start site mutation (H9)	TGTAAAATAATTCAAAATTACCTTTTACAAGCTGTTGCAAGTCGGTG ACATCAGGGTGATCACAGCTTTGAGTCACTGATTAAAAA

Table 3. Primer Sequences Used for Genome Editing		
Primer Name	Letter/Number in Fig.	Sequence
Iso1R	L	CTGCTACCAGGGAAGCAAAA
Iso1SeqF	M	TTCTTTCATGTTGACATCTTTAATTTT
Iso1F	N	GCTTATAATGGCTTGTCTGTTGG
Iso2R	O	TGAAACAATAACCAAATAACATTGG
Iso2SeqR	P	TCTTGATTTGAATCGCAGAAAA
Iso2F	Q	TCAGTAGCCACTATCAAAGACCT
Iso3R	R	TTTTTGAACAATGAATTGGGTTT
Iso3SeqF	S	AGCCTACGCTCAGATCAAGG
Iso3F	T	TTTTTGAACAATGAATTGGGTTT

Chapter 5

Discussion

Summary & Significance

AS is a rare neurodevelopmental disorder for which there is currently no cure, and for which there is no effective treatment.[20] There are, however, multiple promising therapeutic approaches currently being explored. These approaches are based on the use of the human UBE3A sequence: one approach involves unsilencing the paternal copy of UBE3A by targeting the *UBE3A-ATS* transcript[70] while another involves delivery and expression of human UBE3A cDNA via vector-based approaches, such as AAV.[71] Testing these therapies in a mouse model requires the use of the murine sequence or development of a humanized mouse model, which is often expensive. It is therefore necessary for the AS field to develop models using human cells which can be used to test these potential therapies. Of additional importance is establishing robust cellular phenotypes that can be assayed following delivery of these and other potential therapies. Human pluripotent stem cells provide us with a tool to achieve both of these major milestones. One way to improve *in vitro* disease models is through the generation of isogenic pluripotent stem cell lines, which can be differentiated into neurons. Using isogenic cell lines eliminates the natural variation between unrelated individuals, which can in turn cause molecular and phenotypic variation even when comparing neurons derived from individuals considered to be neurotypical.[73,74] Work in this thesis details the generation of isogenic AS and control human pluripotent stem cell (hPSC) lines and establishes the use of these cells and their neuronal derivatives as robust cellular models to study AS.

In chapter 2, I outline the generation of several isogenic iPSC and hESC lines. Some of these lines were used in chapters 3 and 4 of this thesis. Some of these lines have also been used by others to establish other *in vitro* AS phenotypes, such as differences in electrophysiological properties[76] or in neuronal morphology. While these phenotypes can be (and have been) established in non-isogenic neurons, they require significantly more patient-derived lines and more neurons measure from each line to counter natural variation between

the cells. Isogenic AS and control pairs will allow us to measure these same phenotypes using fewer cell lines and/or total numbers of cells overall. Additionally, these isogenic cell lines allow us to ask specific genetic questions of these phenotypes, such as which *in vitro* phenotypes are consistently recapitulated in AS neurons due to loss of *UBE3A*, and not due to the decreased expression of non-imprinted genes that occurs in a large proportion of AS cases (due to large deletions of 15q11-q13).

Two of the isogenic stem cell lines generated in chapter 2 were used in chapter 3 to establish a quantitative molecular AS phenotype by comparing the transcriptomes of AS and control neurons. Using conservative criteria, I found 855 genes that were commonly differentially expressed in each AS line when compared to their respective isogenic control. Importantly, this phenotype was established in cell lines that lacked either UBE3A expression or function, allowing us to establish a list of genes that we should be able to reliably assay in cells derived from any AS patient subtype in the future (large 15q11-q13 deletion, UPD, etc.). I then validated the use of these genes as a phenotype by rescuing some of these gene expression changes after restoring ubiquitin ligase function to a ligase-dead version of UBE3A, which was achieved by the introduction of the E6 oncoprotein to AS NPCs. Based on our findings, I propose that this molecular phenotype, which can be assayed in a manner that is relatively high throughput, can and should be used to assess the efficacy of potential therapies being tested in cultured human AS neurons, particularly those that are meant to restore the expression or function of UBE3A.

In chapter 4, I utilized isogenic stem cell lines that had a mutation at one of the three UBE3A protein isoform translational start sites, creating hESCs that lacked either protein isoform 1, 2, or 3. At the time of writing, there have been few published studies examining the human protein isoforms. Increased understanding of these human protein isoforms is important given that the mouse and human isoforms are not completely conserved. For vector-based gene therapy approaches, knowledge of the isoforms is especially important as each isoform

would potentially need to be delivered in separate vectors if all three need to be replaced in AS patients. Using our isogenic lines, I demonstrated that isoform 1 accounts for the majority of total UBE3A protein in both hESCs and hESC-derived neurons. I also showed that UBE3A predominantly localizes to the cytoplasm in these cell types, that loss of isoform 1 dramatically reduces cytoplasmic protein levels, and that loss of isoforms 2 and 3 does not appear to affect the abundance of UBE3A in any of the subcellular locations examined (cytoplasm, mitochondria, nucleus). Furthermore, I showed that isoform 1-null neurons recapitulate some, but not all, AS phenotypes.

Future Directions

Additional uses for the isogenic stem cell lines

The isogenic lines used to establish the molecular AS phenotype would be useful tools for addressing several other AS questions. One major question in the AS field is: what are UBE3A's targets in neurons? Based on the existence of AS-causing mutations in the UBE3A gene that result in a truncated protein or a ligase-dead protein, it can be assumed that loss of UBE3A's E3 ligase function is a major cause of the phenotypes seen in AS. Efforts to identify and validate UBE3A's targets in neurons have unfortunately not been successful. One potential use for our isogenic cell lines, especially the point mutation and its corrected control, would be to perform proteomic analyses to establish a list of proteins that are increased in the AS compared to the control neurons. Because the point mutation specifically causes loss of UBE3A's ligase function, the proteins identified in this analysis would include accumulated substrates that can no longer be targeted for proteasomal degradation and potentially downstream changes in secondary proteins. Then, to better discern the direct targets of UBE3A, one could acutely knockdown UBE3A in the isogenic control neurons, or any normal neuron, and see which of the targets identified in the first experiment are still affected. Another use for the isogenic lines would be to perform mixing experiments to determine whether *in vitro* AS

phenotypes are cell intrinsic or whether they are a result of a more global neuronal disfunction. For example, one could add a small portion of AS neurons to isogenic controls (such as the AS point mutant and its isogenic control) and determine whether the AS neurons still exhibit impaired synaptic function or morphological phenotypes.

Transcriptome Phenotype

Our main goal in establishing a molecular transcriptome phenotype for AS was to have a phenotype that reflected the loss of UBE3A in AS neurons that could be assayed following restoration of UBE3A's expression or function. The most promising therapeutic approach for AS is to use antisense oligonucleotides designed to target the *UBE3A-ATS* transcript, which, upon its degradation, would reactivate the paternal copy of *UBE3A*.[70] This therapy would be specific to the human genetic sequence, and would therefore need to be tested in either a humanized mouse model, a time consuming and expensive endeavor, or in human neurons. With this transcriptome phenotype, we can now treat any of our AS iPSC-derived neurons with these ASOs and measure the extent to which the transcriptome phenotype is restored upon activation of paternal UBE3A and also determine the effects of the timing of *UBE3A* restoration. While performing mRNAseq would give the most thorough assessment of the degree of phenotype restoration, it is also possible to use a less expensive and less time-consuming approach to assay phenotype restoration in a subset of the DE genes, such as targeted RNAseq or custom qRT-PCR arrays.

In addition to establishing our molecular AS phenotype, it is possible to gain further insight into our AS neurons from our mRNAseq data. Although I used conservative criteria to establish our gene list, gene ontology (GO) analysis of these genes revealed dysregulation of pathways or processes that could be further explored. One striking finding was that the top GO cluster for the downregulated DE genes involved the terms "synaptic signaling", "trans-synaptic signaling", and "chemical synaptic transmission". Interestingly, we also see synaptic phenotypes

in our AS stem cell-derived neurons.[76] It's possible that the subset of synaptic genes highlighted by GO will give us insight into these synaptic deficits. The top GO clusters for upregulated DE genes included terms involved in cell adhesion and the extracellular matrix. These processes could be examined in our cultured AS neurons to see whether they are affected and whether they can be quantified as it could reflect another *in vitro* AS neuron phenotype.

One caveat to our transcriptome phenotype is that it is only a phenotype for forebrain glutamatergic neurons, the major constituent of our cell cultures. However, studies in AS mouse models have indicated that GABAergic neurons likely also have deficits in AS.[112] Therefore, it would be worth establishing a transcriptome phenotype for AS in isogenic GABAergic stem cell-derived neurons. This would provide insight into genes that are commonly differentially expressed in both neuron subtypes, but also tell us which differentially expressed genes are unique to each neuronal subtype.

E6 and UBE3A activation

In addition to validating our transcriptome phenotype, our experiments involving E6 oncoprotein provide an exciting proof of principle that allosteric activation of UBE3A via E6 could potentially restore ligase function to UBE3A in a subset of AS patients. Treatment of our AS NPCs with E6 was able to restore the resting membrane potential phenotype, as well as rescue expression of certain genes. Data from *in vitro* ubiquitination assays, shown here and in other studies, indicate that E6 acts as an allosteric activator of wildtype UBE3A and can further increase its ligase activity.[41] While I am not suggesting that E6 expression itself could serve as a potential AS therapy, it might be possible to identify small molecules or peptides that can similarly serve as allosteric UBE3A activators. It is possible that, in cases where an AS-causing mutation does not cause a complete loss of UBE3A protein, allosteric activation of the remaining UBE3A protein could restore ligase activity to therapeutically useful levels. In the

future, we plan to test this hypothesis by delivering E6 to AS NPCs known to have low levels of UBE3A expression remaining, such as the isoform 1 KO line discussed in chapter 4.

Studying the UBE3A protein isoforms

Although I gained insight into the abundance and localization of the human protein isoforms using our isogenic isoform-null stem cell lines, there are further studies that should be performed to confirm our results. First, it would be useful to generate double isoform knockout lines to confirm our findings that 1) isoform 1 is the predominant isoform, 2) loss of isoforms 2 or 3 do not seem to affect total UBE3A abundance (consistent with isoform 1 being the predominant isoform), and 3) UBE3A predominantly localizes to the cytoplasm, regardless of protein isoform. If our conclusions are correct, I would expect that double isoform-null lines would recapitulate these results by showing relatively normal protein levels in isoform 2/3-null cells and localization of remaining UBE3A to the cytoplasm, by both fractionation and immunostaining, in all isoform double-null combinations. Another question is whether the isoforms have different targets, as all three isoforms are full length and presumably are functional E3 ligases. It might be possible to overexpress individual UBE3A isoforms in UBE3A null cells and examine proteomic changes in the cells.

One major motivation for studying the human protein isoforms is that the knowledge gained could be useful in the development of vector-based gene therapies. This approach involves treating AS patients with a vector containing a UBE3A transgene, which would then allow the patient to express normal UBE3A protein. The vectors being investigated for this approach would likely only be able to express individual cDNAs, which would mean that a vector expressing each isoform would need to be made. This potentially means that three different vectors would have to be used. However, it is also possible that not all three isoforms would need to be delivered to a patient to have a therapeutic effect. One future study would be to treat AS NPCs or neurons with virus expressing only one of the three isoforms to see whether we

could rescue any of our *in vitro* AS phenotypes, or whether a combination of two or more isoforms is necessary to restore these phenotypes.

Conclusions

In this thesis, I have developed several isogenic iPSC and hESC lines that can be used as robust cellular AS models. I have used these isogenic lines to establish a quantitative molecular phenotype for AS neurons that can be assayed following drug treatments. Still other isogenic lines established in this thesis have been used to study the human UBE3A protein isoforms, about which little information is currently known. Finally, I have validated these cell lines and have made them available so that others may also use them to make important AS discoveries. Ultimately, I hope that these lines can be used toward the development of effective therapies for AS.

1. Reis, A. *et al.* Imprinting mutations suggested by abnormal DNA methylation patterns in familial Angelman and Prader-Willi syndromes. *Am. J. Hum. Genet.* **54,** 741–747 (1994).

2. Bailey, J. A. *et al.* Recent segmental duplications in the human genome. *Science* **297,** 1003–1007 (2002).

3. Chamberlain, S. J. & Lalande, M. Neurodevelopmental disorders involving genomic imprinting at human chromosome 15q11-q13. *Neurobiol. Dis.* **39,** 13–20 (2010).

4. Nicholls, R. D., Saitoh, S. & Horsthemke, B. Imprinting in Prader-Willi and Angelman syndromes. *Trends Genet.* **14,** 194–200 (1998).

5. Bartolomei, M. S. & Ferguson-Smith, A. C. Mammalian genomic imprinting. *Cold Spring Harb. Perspect. Biol.* **3,** (2011).

6. Kelsey, G. & Feil, R. New insights into establishment and maintenance of DNA methylation imprints in mammals. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368,** 20110336 (2013).

7. Kanduri, C. Long noncoding RNAs: Lessons from genomic imprinting. *Biochim. Biophys. Acta* **1859,** 102–111 (2016).

8. Plasschaert, R. N. & Bartolomei, M. S. Genomic imprinting in development, growth, behavior and stem cells. *Development* **141,** 1805–1813 (2014).

9. Runte, M. *et al.* The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. *Hum. Mol. Genet.* **10,** 2687–2700 (2001).

10. Chamberlain, S. J. RNAs of the human chromosome 15q11-q13 imprinted region. *Wiley Interdiscip. Rev.* **4,** 155–166 (2013).

11. Rougeulle, C., Glatt, H. & Lalande, M. The Angelman syndrome candidate gene, UBE3A/E6-AP, is imprinted in brain. *Nat. Genet.* **17,** 14–15 (1997).

12. Saitoh, S. *et al.* Minimal definition of the imprinting center and fixation of chromosome 15q11-q13 epigenotype by imprinting mutations. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 7811–7815 (1996).

13. El-Maarri, O. *et al.* Maternal methylation imprints on human chromosome 15 are established during or after fertilization. *Nat. Genet.* **27,** 341–344 (2001).

14. DuBose, A. J., Smith, E. Y., Johnstone, K. A. & Resnick, J. L. Temporal and developmental requirements for the Prader-Willi imprinting center. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 3446–3450 (2012).

15. Kantor, B., Kaufman, Y., Makedonski, K., Razin, A. & Shemer, R. Establishing the epigenetic status of the Prader-Willi/Angelman imprinting center in the gametes and embryo. *Hum. Mol. Genet.* **13,** 2767–2779 (2004).

16. Buiting, K. *et al.* Epimutations in Prader-Willi and Angelman syndromes: a molecular study of 136 patients with an imprinting defect. *Am. J. Hum. Genet.* **72,** 571–577 (2003).

17. Driscoll, D. J., Miller, J. L., Schwartz, S. & Cassidy, S. B. Prader-Willi Syndrome. in (eds. Adam, M. P. et al.) (1993).

18. de Smith, A. J. *et al.* A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs)

is associated with hyperphagia, obesity and hypogonadism. *Hum. Mol. Genet.* **18,** 3257–3265 (2009).

19. Duker, A. L. *et al.* Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader-Willi syndrome. *Eur. J. Hum. Genet.* **18,** 1196–1201 (2010).

20. Dagli, A. I. & Williams, C. A. Angelman Syndrome. in *GeneReviews(R)* (eds. Pagon, R. A. et al.) (University of Washington, Seattle, 1993). doi:NBK1144 [bookaccession]

21. Kishino, T., Lalande, M. & Wagstaff, J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat. Genet.* **15,** 70–73 (1997).

22. Williams, C. A., Driscoll, D. J. & Dagli, A. I. Clinical and genetic aspects of Angelman syndrome. *Genet. Med.* **12,** 385–395 (2010).

23. Finucane, B. M. *et al.* 15q Duplication Syndrome and Related Disorders. in (eds. Adam, M. P. et al.) (1993).

24. Noor, A. *et al.* 15q11.2 Duplication Encompassing Only the UBE3A Gene Is Associated with Developmental Delay and Neuropsychiatric Phenotypes. *Hum. Mutat.* **36,** 689–693 (2015).

25. Scheffner, M., Huibregtse, J. M., Vierstra, R. D. & Howley, P. M. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75,** 495–505 (1993).

26. Scheffner, M. & Kumar, S. Mammalian HECT ubiquitin-protein ligases: biological and pathophysiological aspects. *Biochim. Biophys. Acta* **1843,** 61–74 (2014).

27. Cooper, E. M., Hudson, A. W., Amos, J., Wagstaff, J. & Howley, P. M. Biochemical analysis of Angelman syndrome-associated mutations in the E3 ubiquitin ligase E6-associated protein. *J. Biol. Chem.* **279,** 41208–41217 (2004).

28. Tomaic, V. & Banks, L. Angelman syndrome-associated ubiquitin ligase UBE3A/E6AP mutants interfere with the proteolytic activity of the proteasome. *Cell Death Dis.* **6,** e1625 (2015).

29. Nuber, U., Schwarz, S. E. & Scheffner, M. The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as its own substrate. *Eur. J. Biochem.* **254,** 643–649 (1998).

30. Talis, A. L., Huibregtse, J. M. & Howley, P. M. The role of E6AP in the regulation of p53 protein levels in human papillomavirus (HPV)-positive and HPV-negative cells. *J. Biol. Chem.* **273,** 6439–6445 (1998).

31. Kumar, S., Talis, A. L. & Howley, P. M. Identification of HHR23A as a substrate for E6-associated protein-mediated ubiquitination. *J. Biol. Chem.* **274,** 18785–18792 (1999).

32. Margolis, S. S. *et al.* EphB-mediated degradation of the RhoA GEF Ephexin5 relieves a developmental brake on excitatory synapse formation. *Cell* **143,** 442–455 (2010).

33. Zaaroor-Regev, D. *et al.* Regulation of the polycomb protein Ring1B by self-ubiquitination or by E6-AP may have implications to the pathogenesis of Angelman syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 6788–6793 (2010).

34. Reiter, L. T., Seagroves, T. N., Bowers, M. & Bier, E. Expression of the Rho-GEF Pbl/ECT2 is regulated by the UBE3A E3 ubiquitin ligase. *Hum. Mol. Genet.* **15,** 2825–

2835 (2006).

35. Martinez-Noel, G. *et al.* Identification and proteomic analysis of distinct UBE3A/E6AP protein complexes. *Mol. Cell. Biol.* **32,** 3095–3106 (2012).

36. Kuhnle, S., Mothes, B., Matentzoglu, K. & Scheffner, M. Role of the ubiquitin ligase E6AP/UBE3A in controlling levels of the synaptic protein Arc. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 8888–8893 (2013).

37. Kuhne, C. & Banks, L. E3-ubiquitin ligase/E6-AP links multicopy maintenance protein 7 to the ubiquitination pathway by a novel motif, the L2G box. *J. Biol. Chem.* **273,** 34302–34309 (1998).

38. Li, L., Li, Z., Howley, P. M. & Sacks, D. B. E6AP and calmodulin reciprocally regulate estrogen receptor stability. *J. Biol. Chem.* **281,** 1978–1985 (2006).

39. Louria-Hayon, I. *et al.* E6AP promotes the degradation of the PML tumor suppressor. *Cell Death Differ.* **16,** 1156–1166 (2009).

40. Nawaz, Z. *et al.* The Angelman syndrome-associated protein, E6-AP, is a coactivator for the nuclear hormone receptor superfamily. *Mol. Cell. Biol.* **19,** 1182–1189 (1999).

41. Mortensen, F. *et al.* Role of ubiquitin and the HPV E6 oncoprotein in E6AP-mediated ubiquitination. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 9872–9877 (2015).

42. Kuhnle, S. *et al.* Physical and functional interaction of the HECT ubiquitin-protein ligases E6AP and HERC2. *J. Biol. Chem.* **286,** 19410–19416 (2011).

43. Kleijnen, M. F. *et al.* The hPLIC proteins may provide a link between the ubiquitination machinery and the proteasome. *Mol. Cell* **6,** 409–419 (2000).

44. Smith, C. L. *et al.* Genetic ablation of the steroid receptor coactivator-ubiquitin ligase, E6-AP, results in tissue-selective steroid hormone resistance and defects in reproduction. *Mol. Cell. Biol.* **22,** 525–535 (2002).

45. Singhmar, P. & Kumar, A. Angelman syndrome protein UBE3A interacts with primary microcephaly protein ASPM, localizes to centrosomes and regulates chromosome segregation. *PLoS One* **6,** e20397 (2011).

46. Yamamoto, Y., Huibregtse, J. M. & Howley, P. M. The human E6-AP gene (UBE3A) encodes three potential protein isoforms generated by differential splicing. *Genomics* **41,** 263–266 (1997).

47. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).

48. Aken, B. L. *et al.* Ensembl 2017. *Nucleic Acids Res.* **45,** D635–D642 (2017).

49. Miao, S. *et al.* The Angelman syndrome protein Ube3a is required for polarized dendrite morphogenesis in pyramidal neurons. *J. Neurosci.* **33,** 327–333 (2013).

50. Valluy, J. *et al.* A coding-independent function of an alternative Ube3a transcript during neuronal development. *Nat. Neurosci.* **18,** 666–673 (2015).

51. Trickett, J., Heald, M. & Oliver, C. Sleep in children with Angelman syndrome: Parental concerns and priorities. *Res. Dev. Disabil.* **69,** 105–115 (2017).

52. Prasad, A., Grocott, O., Parkin, K., Larson, A. & Thibert, R. L. Angelman syndrome in adolescence and adulthood: A retrospective chart review of 53 cases. *Am. J. Med. Genet. A* **176,** 1327–1334 (2018).

53. Forrest, K. M., Young, H., Dale, R. C. & Gill, D. S. Benefit of corticosteroid therapy in Angelman syndrome. *J. Child Neurol.* **24,** 952–958 (2009).

54. Grocott, O. R., Herrington, K. S., Pfeifer, H. H., Thiele, E. A. & Thibert, R. L. Low glycemic index treatment for seizure control in Angelman syndrome: A case series from the Center for Dietary Therapy of Epilepsy at the Massachusetts General Hospital. *Epilepsy Behav.* **68,** 45–50 (2017).

55. Grieco, J. C. *et al.* An open-label pilot trial of minocycline in children as a treatment for Angelman syndrome. *BMC Neurol.* **14,** 232 (2014).

56. Tan, W.-H. *et al.* A randomized controlled trial of levodopa in patients with Angelman syndrome. *Am. J. Med. Genet. A* **176,** 1099–1107 (2018).

57. Ovid Therapeutics, I. Ovid Therapeutics Announces Positive Topline Data from Phase 2 STARS Trial of OV101 for the Treatment of Angelman Syndrome. *[Press Release]* (2018).

58. Silva-Santos, S. *et al.* Ube3a reinstatement identifies distinct developmental windows in a murine Angelman syndrome model. *J. Clin. Invest.* **125,** 2069–2076 (2015).

59. Rotaru, D. C., van Woerden, G. M., Wallaard, I. & Elgersma, Y. Adult Ube3a Gene Reinstatement Restores the Electrophysiological Deficits of Prefrontal Cortex Layer 5 Neurons in a Mouse Model of Angelman Syndrome. *J. Neurosci.* **38,** 8011–8030 (2018).

60. Meng, L. *et al.* Truncation of Ube3a-ATS unsilences paternal Ube3a and ameliorates behavioral defects in the Angelman syndrome mouse model. *PLoS Genet.* **9,** e1004039 (2013).

61. Jacob, R. A. *et al.* Moderate folate depletion increases plasma homocysteine and decreases lymphocyte DNA methylation in postmenopausal women. *J. Nutr.* **128,** 1204–1212 (1998).

62. Cooney, C. A., Dave, A. A. & Wolff, G. L. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J. Nutr.* **132,** 2393S–2400S (2002).

63. Peters, S. U. *et al.* Double-blind therapeutic trial in Angelman syndrome using betaine and folic acid. *Am. J. Med. Genet. A* **152A,** 1994–2001 (2010).

64. Bird, L. M. *et al.* A therapeutic trial of pro-methylation dietary supplements in Angelman syndrome. *Am. J. Med. Genet. A* **155A,** 2956–2963 (2011).

65. Huang, H. S. *et al.* Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons. *Nature* **481,** 185–189 (2011).

66. Mabb, A. M. *et al.* Topoisomerase 1 Regulates Gene Expression in Neurons through Cleavage Complex-Dependent and -Independent Mechanisms. *PLoS One* **11,** e0156439 (2016).

67. King, I. F. *et al.* Topoisomerases facilitate transcription of long genes linked to autism. *Nature* **501,** 58–62 (2013).

68. Mabb, A. M. *et al.* Topoisomerase 1 inhibition reversibly impairs synaptic function. *Proc.*

Natl. Acad. Sci. U. S. A. **111,** 17290–17295 (2014).

69. Lee, H.-M. *et al.* Characterization and structure-activity relationships of indenoisoquinoline-derived topoisomerase I inhibitors in unsilencing the dormant Ube3a gene associated with Angelman syndrome. *Mol. Autism* **9,** 45 (2018).

70. Meng, L. *et al.* Towards a therapy for Angelman syndrome by targeting a long non-coding RNA. *Nature* (2014). doi:10.1038/nature13975 [doi]

71. Daily, J. L. *et al.* Adeno-associated virus-mediated rescue of the cognitive defects in a mouse model for Angelman syndrome. *PLoS One* **6,** e27221 (2011).

72. PTC Biotherapeutics. PTC Biotherapeutics Pipeline. (2018). Available at: https://www.ptcbio.com/en/pipeline/.

73. Pickrell, J. K. *et al.* Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* **464,** 768–772 (2010).

74. Storey, J. D. *et al.* Gene-expression variation within and among human populations. *Am. J. Hum. Genet.* **80,** 502–509 (2007).

75. Landers, M. *et al.* Regulation of the large (approximately 1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn. *Nucleic Acids Res.* **32,** 3480–3492 (2004).

76. Fink, J. J. *et al.* Disrupted neuronal maturation in Angelman syndrome-derived induced pluripotent stem cells. *Nat. Commun.* **8,** (2017).

77. Sadhwani, A. *et al.* Two Angelman families with unusually advanced neurodevelopment carry a start codon variant in the most highly expressed UBE3A isoform. *Am. J. Med. Genet. A* **176,** 1641–1647 (2018).

78. Chamberlain, S. J. *et al.* Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader-Willi syndromes. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 17668–17673 (2010).

79. Germain, N. D. *et al.* Gene expression analysis of human induced pluripotent stem cell-derived neurons carrying copy number variants of chromosome 15q11-q13.1. *Mol. Autism* **5,** 44 (2014).

80. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nature methods* **11,** 783–784 (2014).

81. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819–823 (2013).

82. Kato-Inui, T., Takahashi, G., Hsu, S. & Miyaoka, Y. Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 with improved proof-reading enhances homology-directed repair. *Nucleic Acids Res.* **46,** 4677–4688 (2018).

83. Yang, L. *et al.* Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* **41,** 9049–9061 (2013).

84. Yu, C. *et al.* Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16,** 142–147 (2015).

85. Li, G. *et al.* Small molecules enhance CRISPR/Cas9-mediated homology-directed

genome editing in primary cells. *Sci. Rep.* **7,** 8943 (2017).

86. Yachdav, G. *et al.* PredictProtein - An open resource for online prediction of protein structural and functional features. *Nucleic Acids Res.* **42,** 337–343 (2014).

87. Bendl, J. *et al.* PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *PLoS Comput. Biol.* **10,** 1–11 (2014).

88. Kraft, K. *et al.* Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. *Cell Rep.* (2015). doi:10.1016/j.celrep.2015.01.016

89. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42,** e168 (2014).

90. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533,** 125–129 (2016).

91. Zhang, J.-P. *et al.* Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* **18,** 35 (2017).

92. Somers, A. *et al.* Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. *Stem Cells* **28,** 1728–1740 (2010).

93. Martins-Taylor, K. *et al.* Imprinted expression of UBE3A in non-neuronal cells from a Prader-Willi syndrome patient with an atypical deletion. *Hum. Mol. Genet.* **23,** 2364–2373 (2014).

94. Germain, N. D., Banda, E. C., Becker, S., Naegele, J. R. & Grabel, L. B. Derivation and isolation of NKX2.1-positive basal forebrain progenitors from human embryonic stem cells. *Stem Cells Dev.* **22,** 1477–1489 (2013).

95. Chen, P.-F., Hsiao, J. S., Sirois, C. L. & Chamberlain, S. J. RBFOX1 and RBFOX2 are dispensable in iPSCs and iPSC-derived neurons and do not contribute to neural-specific paternal UBE3A silencing. *Sci. Rep.* **6,** (2016).

96. Langouet, M. *et al.* Zinc finger protein 274 regulates imprinted expression of transcripts in Prader-Willi syndrome neurons. *Hum. Mol. Genet.* **27,** 505–515 (2018).

97. Chamberlain, S. J., Germain, N. D., Chen, P.-F., Hsiao, J. S. & Glatt-Deeley, H. Modeling Genomic Imprinting Disorders Using Induced Pluripotent Stem Cells. *Methods Mol. Biol.* **1353,** 45–64 (2016).

98. Leek, J. T. svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res.* **42,** (2014).

99. Huibregtse, J. M., Scheffner, M. & Howley, P. M. Cloning and expression of the cDNA for E6-AP, a protein that mediates the interaction of the human papillomavirus E6 oncoprotein with p53. *Mol. Cell. Biol.* **13,** 775–784 (1993).

100. Lois, C., Hong, E. J., Pease, S., Brown, E. J. & Baltimore, D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295,** 868–872 (2002).

101. Kuballa, P., Matentzoglu, K. & Scheffner, M. The role of the ubiquitin ligase E6-AP in human papillomavirus E6-mediated degradation of PDZ domain-containing proteins. *J.*

Biol. Chem. **282,** 65–71 (2007).

102. Volpato, V. *et al.* Reproducibility of Molecular Phenotypes after Long-Term Differentiation to Human iPSC-Derived Neurons: A Multi-Site Omics Study. *Stem cell reports* (2018). doi:10.1016/j.stemcr.2018.08.013

103. Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* **32,** 896–902 (2014).

104. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30,** 2114–2120 (2014).

105. Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32,** 3047–3048 (2016).

106. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).

107. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

108. Fresno, C. & Fernandez, E. A. RDAVIDWebService: a versatile R interface to DAVID. *Bioinformatics* **29,** 2810–2811 (2013).

109. Cotney, J. L. & Noonan, J. P. Chromatin immunoprecipitation with fixed animal tissues and preparation for high-throughput sequencing. *Cold Spring Harb. Protoc.* **2015,** 191–199 (2015).

110. Wilderman, A., VanOudenhove, J., Kron, J., Noonan, J. P. & Cotney, J. High-Resolution Epigenomic Atlas of Human Embryonic Craniofacial Development. *Cell Rep.* **23,** 1581–1597 (2018).

111. Judson, M. C., Sosa-Pagan, J. O., Del Cid, W. A., Han, J. E. & Philpot, B. D. Allelic specificity of Ube3a expression in the mouse brain during postnatal Development. *J. Comp. Neurol.* **522,** 1874–1896 (2014).

112. Judson, M. C. *et al.* GABAergic Neuron-Specific Loss of Ube3a Causes Angelman Syndrome-Like EEG Abnormalities and Enhances Seizure Susceptibility. *Neuron* **90,** 56–69 (2016).